THREE ESSAYS ON FINANCIAL ECONOMICS

BY

LIFENG GU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Finance
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Doctoral Committee:

Associate Professor Dirk Hackbarth, Chair
Professor Timothy C Johnson
Professor George Pennacchi
Assistant Professor Prachi Deuskar

Abstract

My first essay, *Product Market Competition, R&D Investment and Stock Returns*, considers the interaction between product market competition and investment in research and development (R&D) to tackle two asset pricing puzzles: the positive R&D-return relation and the positive competition-return relation. Using a standard model of R&D return dynamics, I establish that competition and R&D investments have a strong interaction effect on stock returns. It is more likely that firms with high R&D expenditures will end up with very low returns on their ventures because rival firms win the innovation race. Because there are more potential rival firms in competitive industries, R&D intensive firms in competitive industries are riskier. Consistent with the predictions of the model, I find a robust empirical relation between R&D intensity and stock returns, but only in competitive industries. This finding suggests that the risk derived from product market competition has important asset pricing implications and potentially drives a large portion of the positive R&D-return relation. Furthermore, firms in competitive industries earn higher returns than firms in concentrated industries only among R&D-intensive firms. My second finding therefore provides a risk-based explanation for the heretofore puzzling competition premium.

The second essay, *Governance and Equity Prices: Does Transparency Matter?*, examines how a firm's information environment and corporate governance interact. Firms with higher takeover vulnerability are associated with higher abnormal returns, but even more so if they also have higher accounting transparency. The effect is largely monotonic — it is small and insignificant for opaque firms and large and significant for transparent firms — and it holds in sample splits based on firm characteristics, such as leverage or size. A portfolio that buys firms with the highest level of takeover vulnerability and shorts firms with the lowest level of takeover vulnerability — provided it includes the tercile of transparent firms as defined by the first principal component of forecast error, forecast dispersion, and revision volatility — generates a monthly abnormal return of 1.37% for value-weighted (1.28% for equal-weighted) portfolios, which is nearly twice as large as the alpha in the full sample. This result survives numerous robustness tests, and it also remains large and significant even if we extend the sample period until 2006. Hence transparency and governance (i.e., takeover vulnerability) are complements. This complementarity effect is consistent with the view that more transparent firms are more likely to be taken over, since acquirers can bid more effectively and identify synergies more precisely.

The third essay, *Takeover Likelihood, Firm Transparency and the Cross-Section of Returns*, explores the determinants of a firm's takeover likelihood and proposes to include the firm's in-

formation environment as additional predicting variable since a transparent environment could facilitate takeovers by making it easier for bidders to value the firm and the synergy of the deal. The logit estimation including this new variable over the sample period of 1991 to 2009 produces results consistent with this view and better fits the real takeover data. The new takeover factor constructed as the return to the long-short portfolio that buys firms with top takeover probability and sells firms with bottom takeover probability better captures the variation in the cross-section of stock returns.

To Father and Mother.

Acknowledgments

I am very grateful to my adviser, Dirk Hackbarth, for his encouragement and guidance. Many thanks to my committee members, Timothy Johnson, George Pennacchi, Prachi Deuskar for their advise and support. Also I would like to thank all other members in the Department of Finance at the University of Illinois at Urbana-Champaign for their assistance. Finally, special thanks to my parents who have been providing me with love and support during this long journey.

Table of Contents

Chapter 1

Product Market Competition, R&D Investment and Stock Returns

1.1 Introduction

Investment in research and development (R&D) is one of the crucial elements of corporate activities that drive companies' long-term viability. A large share of the US stock market consists of firms that invest in R&D intensively. Often a firm enters into an innovation race with many rivals. If a competitor successfully completes the project first, then other firms may have to suspend or even abandon their projects. Because R&D investment tends to be irreversible, suspension or abandonment of the R&D project significantly reduces firm value. Therefore, competition can have a large impact on R&D-intensive firms.

Prior research in asset pricing has uncovered two asset pricing puzzles: the positive R&D-return relation and the positive competition-return relation.[1] However, they have largely been regarded as unrelated to each other, and we still lack a thorough understanding of the real mechanism driving these two return patterns. Thus research on the interaction effect between R&D investment and competition can shed new light on these two puzzles.

This paper studies the joint effect of product market competition and R&D investment on stock returns. My analysis is based on a partial equilibrium model for a multi-stage R&D venture by Berk, Green, and Naik (2004). Specifically, I establish two hypotheses: (1) The positive R&D-return relation is stronger in competitive industries; (2) The positive competition-return relation strengthens among R&D-intensive firms. In other words, there is a strong positive interaction effect

[1] The two asset pricing puzzles are the positive R&D-return relation (Chan, Lakonishok, and Sougiannis (2001)) and the competition premium (Hou and Robinson (2006)). In the literature of R&D investments and stock returns, researchers have proposed different ways to measure firms' innovation activity such as R&D investments, R&D increment, R&D efficiency and R&D ability and consistently find positive premiums associated with those measures. However, we still lack a thorough understanding of the real mechanism driving this return pattern. Hou and Robinson (2006) show that firms in competitive industries outperform firms in concentrated industries. They speculate that this might be because either barriers to enter concentrated industries insulate firms from undiversifiable distress risk or concentrated industries are less innovative. However, so far no article has provided any formal tests or explanation to address the mechanism underlying the competition premium.

between competition and R&D investment on stock returns.

In the model, the firm progresses through the R&D project stage by stage and decides whether to incur an instantaneous R&D investment to continue the project or not. Prior to completion, the decision maker can observe the future cash flows that the project would be producing if it were completed today. Since the risk associated with cash flows has a systematic component, this feature imparts to the project a large amount of systematic risk, and the R&D venture can be considered as a series of compound options on systematic uncertainty. Since options have higher systematic risk than the underlying asset because of implicit leverage, the R&D venture demands a higher risk premium than the stochastic cash flow itself.

Competition increases the probability that the potential future cash flows will be extinguished, and thus decreases the benefit of investing and raises the chance of project suspension in the event of an adverse shock to the future cash flow. Therefore, firms' investment decisions and value are more sensitive to the systematic risk that the cash flow carries. Moreover, this negative impact of a potential extinguishing of cash flow is more pronounced for firms with high R&D inputs because high R&D inputs further reduce the value of the option to invest. As a result, the model implies a strong positive relation between competition and stock returns among R&D-intensive firms. It also predicts a stronger positive relation between R&D investment and returns for firms surrounded by high competition.

I test the model's predictions empirically with firm-level data over the 1963 to 2009 period and document a robust positive interaction effect between product market competition and R&D investment on stock returns. Following the literature, I employ a Herfindahl (concentration) index to gauge the competitiveness of an industry, and I use four alternative measures to proxy for firm's R&D intensity: R&D expenditure scaled by net sales, R&D expenditure scaled by total assets, R&D expenditure scaled by capital expenditure, and R&D capital scaled by total assets. The purpose of using different measures is to make sure that the results are not driven by the size effect. Following Chan, Lakonishok, and Sougiannis (2001), I compute R&D capital as the sum of the depreciated R&D expenditures over the past five years, assuming an annual depreciation rate of 20%.

First, I adopt the portfolio sorting approach to test the model's predictions. Specifically, in June of each year, firms are sorted independently into three groups (bottom 30%, middle 40%, top 30%) based on the Herfindahl index and R&D intensity in the previous year.[2] The abnormal returns to the resulting nine portfolios are computed adjusting for the Fama-French three-factor model and the Carhart (1997) four-factor model.

My first finding is that the positive R&D-return relation only exists in competitive industries. More precisely, the abnormal return to the double-sorted portfolio increases monotonically with R&D intensity for firms in competitive industries, but this return pattern does not exist for firms from concentrated industries. This finding holds for all four measures of R&D intensity, and for both equal-weighted and value-weighted portfolio returns. For instance, when R&D intensity is measured

[2]Sorting on these two measures sequentially produces qualitatively and quantitatively similar or even stronger results.

by R&D capital scaled by assets, in competitive industries, the monthly equal-weighted four-factor abnormal returns for the low, medium, and high R&D intensity portfolios are 0.06%, 0.38%, and 0.78%, with t-statistics of 0.72, 4.72 and 4.92, respectively. This results in a 72 basis points monthly return difference between the high and low R&D intensity portfolios. In contrast, for concentrated industries, the abnormal returns for the low, medium, and high R&D intensity portfolios are much smaller and insignificant: they are 0.09%, 0.31%, and 0.22%, with t-statistics of 0.42, 1.60 and 0.41, respectively. This translates to a monthly return spread of 0.13% with a t-statistic of 0.05.

My second finding is that the positive competition-return relation only exists among R&D intensive firms. More specifically, the portfolio abnormal return increases monotonically with industry competition level for firms with high R&D inputs, but this pattern does not exist for firms with low R&D inputs. For example, when R&D intensity is measured by R&D expenditure scaled by sales, for R&D-intensive firms, the monthly equal-weighted four-factor abnormal returns for the low, medium, and high competition portfolios are –0.21%, 0.42%, and 0.63%, with t-statistics of 0.54, 1.90 and 4.08, respectively. The monthly return difference between the high and low competition portfolio is 0.84% with a statistical significance at 5% level. In contrast, for firms with low R&D inputs, the abnormal returns of the competition portfolios are small and insignificant. The monthly abnormal return spread between the high and low competition portfolio is –0.22% and insignificant.

The positive interaction effect between R&D intensity and product market competition is also confirmed by the estimation results of Fama-MacBeth cross-sectional return regressions. Specifically, the coefficient on the interaction term, $R\&D * HHI$, is negative and significant at the 5% level for all measures of R&D intensity.[3] For example, when R&D intensity is proxyed by R&D expenditure scaled by sales, the slope on the interaction term is –1.15% and significant at the 5% level. It is –1.08% and significant at the 5% level when R&D is measured by R&D capital scaled by assets. Moreover, I perform the same tests using dummy variables instead of continuous variables for R&D intensity and industry competition. I find that the coefficient on the interaction term for R&D-intensive firms in competitive industries, $R\&D_{high} * HHI_{low}$, is always positive and significant, while the coefficient on other interaction terms is much smaller and insignificant in almost all cases. For example, when R&D intensity is measured by R&D expenditure scaled by sales, the coefficient on the interaction term, $R\&D_{high} * HHI_{low}$, is 0.42% and significant at the 1% level. This is consistent with the portfolio sorting results that there is always a positive and significant abnormal return associated with the portfolio of high R&D intensity and high competition.

My main findings survive numerous robustness tests. I re-estimate the α using alternative asset pricing models which are extensions of the Carhart (1997) four-factor model with additional factors proposed in the literature, including the liquidity factor of Pastor and Stambaugh (2003); the takeover factor of Cremers, John, and Nair (2009); and the misvaluation factor of Hirshleifer and Jiang (2010). Interestingly, none of these factors weaken the abnormal returns, indicating that liquidity, takeover probability and investor misvaluation are not the reasons driving the main

[3]Since high HHI index value means low competiton level, a negative sign shows a positive interaction effect between R&D intensity and industry competition.

findings documented in this article.

In addition to these tests, I also experiment with dividing the full sample into financial-constrained and financial-unconstrained subsamples or into high-innovation-ability or low-innovation-ability subsamples to verify that the main result is not due to these two firm characteristics that are identified to affect R&D investment's risk or effectiveness. Li (2011) finds that the positive R&D-return relation is driven by financial constraints. She documents that the positive R&D-return relation exists only among financially constrained firms. I examine whether my results are driven by financial constraints by dividing the full sample into financially-constrained and financially-unconstrained subsamples. I find that the positive R&D-return relation is present for both financially constrained and unconstrained firms, only if they are also from competitive industries. This finding indicates that financial constraints don't seem to affect the R&D-return relation once I control for competition. Cohen, Diether, and Malloy (2011) provide evidence that R&D predicts future returns only when firms have high ability to translate the outcome of those innovation projects into real sales growth. They document a significantly positive adjusted return associated with *GOODR&D* (i.e., high R&D intensity and high ability) firms. I test whether my results are driven by innovation ability by dividing the full sample into high-innovation-ability or low-innovation-ability subsamples. The positive R&D-return relation exists for both subsamples, but only when the firms are also from competitive industries. This result leads to another interesting finding that innovation ability does not seem to affect the R&D-return relation once I control for competition. Overall, the main findings along with the results of these robustness tests show that product market competition creates another important dimension that matters for the link between R&D investment and stock returns.

I also investigate whether investor limited attention can explain my findings by dividing the whole sample into subsamples based on firm size, firm analyst coverage, and firm idiosyncratic volatility. The fact that my findings exist in all subsamples indicate that investor limited attention is not the main driver. I further provide evidence in supportive of the risk hypothesis by showing that the portfolio of R&D intensive firms in competitive industries are actually associated with higher cash flow risk by examining the volatility of firm performance measures such as return on assets, return on equity, and profit margin. Thus the test results indicate that the positive and significant abnormal returns are driven by risk factors that are not captured by existing asset pricing models.

To further explore the asset pricing implications of firm innovation, I construct a new factor, the innovation factor, following the procedure in Fama-French (1993) and Chen, Novy-Marx and Zhang (2011). The innovation factor is formed as the weighted-average return of a zero-cost trading strategy that takes a long position in stocks with high innovation levels and a short position in stocks with low innovation levels. This strategy earns a significant average return of 37 basis points per month over the sample period.

Furthermore, I test the performance of this innovation factor using the competition-minus-concentration portfolio in Hou and Robinson (2006) who document that firms in competitive industries earn higher returns than firms in concentrated industries and this competition premium persists after adjusting for common risk factors. Interestingly, augmenting the Carhart (1997) four-factor model with the innovation factor brings the abnormal return of the competition-minus-concentration hedge portfolio down to an insignificant level with a much smaller magnitude. This finding therefore provides a risk-based explanation for the competition premium.

This article has two main contributions. First, it contributes to the literature on the relation between R&D investment and stock returns (e.g., Chan, Lakonishok, and Sougiannis 2001; Chambers, Jennings, and Thompson 2002; Lin 2007; Li 2011; Cohen, Diether, and Malloy 2012). I show that the positive R&D-return relation exists only for firms from competitive industries. This empirical finding along with the implications developed from the model indicate that competition has a large impact on the risk and return profiles of R&D-intensive firms and it potentially drives a large portion of the positive R&D-return relation. Second, this article also contributes to the research on the relation between competition and stock returns. I document a robust empirical relation between competition and return, but only among firms with intensive R&D inputs. This suggests that firms engaging in high levels of innovation activities in competitive industries are actually contributing to this positive competition premium. Thus the finding in this article also provides a risk-based explanation for the heretofore puzzling competition-return relation documented in Hou and Robinson (2006).

The rest of the paper is organized as follows. Section 2 contains the development of testable hypotheses. Section 3 provides details of the data sources, sample selection, variable definitions. Section 4 examines the joint effect of product market competition and R&D investments on stock returns empirically. Main results, robustness checks and interpretations are presented in this section. Finally, Section 5 concludes.

1.2 Hypotheses development

To study the interaction effect of industry competition and R&D investment on risk premium, I adopt the model from Berk, Green, and Naik (2004), who develop a partial equilibrium model for a single multi-stage R&D venture. I provide a brief overview of the model before I derive the two testable hypotheses that I will test with the data in the empirical part of the paper.

1.2.1 Overview of the model

The firm operates in continuous time and progresses through the R&D project stage by stage. It receives a stream of stochastic cash flows after it successfully completes N discrete stages. At any point in time prior to completion, the manager has to make investment decisions to maximize the firm's intrinsic value. Before proceeding to the valuation details of the R&D venture, I discuss several important model assumptions:

1. An important feature of the model is that it assumes the decision maker can observe the future cash flow the project would be producing were it complete today and the decision to continue the project is made based on this information. Therefore the systematic risk of the future cash

flow is imparted to firm's investment decisions. Thus the R&D project can be considered as a series of compound options on the underlying cash flows which have systematic uncertainty. Since options have higher systematic risk than the underlying asset because of implicit leverage, the R&D venture demands a higher risk premium than the stochastic cash flow itself. Take gold mining as an concrete example. The successes and failures in gold mining sites will average out when the company has thousands of mining sites in different locations. Often this type of "geological risks" are taken as diversifiable in textbooks. However, the decision to continue mining at any particular site is made with the current price of gold in mind.

2. Success intensity $\pi(n(t))$: The number of completed stages at time t is denoted by $n(t)$. At any point in time prior to completion, the firm must make R&D investment decisions. Conditional on investment, the probability that the firm will complete the current stage and advance to the next stage in the next instant is πdt. π is the success intensity. In the simplified case in Berk, Green, and Naik (2004)'s model, π is a known function of $n(t)$. For simplicity, I assume it as a constant over all stages. Thus, I write it as π hereafter.

3. Investment cost $RD(n(t))$: If the firm decides to continue the project, it has to incur an investment cost. The R&D cost can be a known function of the number of completed stages $n(t)$. For simplicity, I set it as a constant over all stages. Thus, I write it as RD hereafter. Note that the investment level is not a choice variable.

4. Competition or obsolescence risk: It is assumed that the fixed probability that the potential or actual cash flow is extinguished over the next instant is ϕdt. ϕ is the obsolescence rate. In the model, this risk is idiosyncratic and does not demand any risk premium itself. However, a high probability of obsolescence indicates that it is more likely that the potential cash flows will become zero. Thus a high obsolescence rate lowers the benefit of investing and the value of the option to continue the project. Therefore, it affects the firm's decision to continue or suspend the project and thus indirectly its systematic risk and risk premium.

1.2.2 Valuation

After the firm successfully completes N discrete stages, it receives a stream of stochastic cash flow, $c(t)$, which follows a geometric Brownian motion:

$$dc(t) = \mu c(t)dt + \sigma c(t)dw(t), \tag{1.1}$$

where μ is the constant growth rate of cash flows, σ is the constant standard deviation of cash flows, and $dw(t)$ is an increment of a Wiener process.

The partial equilibrium model adopts an exogenous pricing kernel, which is given by the following process

$$dm(t) = -rm(t)dt - \theta m(t)dz(t), \tag{1.2}$$

where r is the constant risk-free rate and the risk premium for the cash flow process $c(t)$ is denoted as

$$\lambda = \sigma\theta\rho, \tag{1.3}$$

where ρ is the correlation between the Brownian motion processes $w(t)$ and $z(t)$.

Under the risk-neutral measure, the cash flow process $c(t)$ follows a geometric Brownian motion:

$$dc(t) = \hat{\mu}c(t)dt + \sigma c(t)d\hat{w}(t), \tag{1.4}$$

where $\hat{w}(t)$ is a Brownian motion under the risk-neutral measure and $\hat{\mu} = \mu - \lambda$ is the constant drift term of the cash flow process under the risk-neutral measure.

Upon making decisions about whether to continue investing or not, the firm's manager observes (1) the number of completed stages, $n(t)$; (2) the level of cash flow the project would be producing were the project complete already; and (3) whether the firm's potential cash flows have been extinguished through obsolescence. Note that both success intensity π and investment cost RD are assumed to be known functions of $n(t)$. If the cash flow is extinguished, the firm value becomes zero. Thus conditional on the project being alive, the firm value at time t depends on the future cash flow, $c(t)$, and the number of completed stages, $n(t)$. It is denoted by $V(c(t), n(t))$. For simplicity, I write it as $V(c, n)$ hereafter.

If the project is completed successfully, the firm receives a stream of stochastic cash flows. Thus at that point the firm value is simply the continuous-time version of the Gordon-Williams growth model. At any time t prior to the completion of the project, the firm value is the maximum of the firm's risk-neutral expected discounted future value at time T (T is an arbitrary point in the future) plus the discounted value of any cash flows received from time t to T. This multi-stage investment problem is given as

$$V(c, n) = \max_{u(s)\in\{0,1\}, s\in(t,T)} E_t^Q\{e^{-(r+\phi)(T-t)}V(c(T), n(T)) + \int_t^T e^{-(r+\phi)(s-t)}(\nu(s)c(s) - u(s)RD)ds\}, \tag{1.5}$$

where u denotes the decision variable with $u = 1$ if the firm continues investing over the next instant and 0 otherwise, and $E^Q[\,]$ is the expectation operator under the risk-neutral (pricing) measure, Q. ν is an indicator variable, taking value of 0 or 1 to keep track of whether the project is completed or not ($\nu = 1$ indicates that the project is completed and $n(t) = N$. $\nu = 0$ indicates that the project is not completed and $n(t) < N$.). RD is the instantaneous R&D cost if the firm continues investing over the next instant. Note that the level of the R&D investment is not a choice variable.

ϕ is the obsolescence rate; that is, the probability of the occurrence of an obsolescence event which could result in zero future cash flows in the next instant. In competitive industries, many firms are competing in the development of a new product or technology. Once one firm claims victory, it will take the cash flows produced by the new product and other firms face a project with zero future cash flows and thus have to restructure or abandon the R&D project. In contrast, in non-competitive industries, there are fewer competitors and firms can conduct the development with less fear about the breakthrough of rivals. Therefore, ϕ will tend to be higher in competitive industries.

By applying Ito's lemma to the value function $V(c,n)$ we can derive the Hamilton-Bellman-Jacobi equation of the investment problem. It is given as follows:

$$(r+\phi)V(c,n) = \frac{1}{2}\sigma^2 c^2 \frac{\partial^2}{\partial c^2}V(c,n) + \hat{\mu} c \frac{\partial}{\partial c}V(c,n) + \max_{u\in\{0,1\}} u\pi[V(c,n+1) - V(c,n)] - RD, \quad (1.6)$$

where π denotes the probability that the firm will successfully complete the current stage over the next instant if the firm invests. Once it successfully advances to the next stage, the firm value will jump to $V(c,n+1)$. Thus, $\pi[V(c,n+1) - V(c,n)]$ is the expected jump in firm value after the investment or the benefit of investing. Consider a biotechnology firm as an example. Its value usually jumps after the firm advances to more advanced phases.

At each instant of time, the firm makes an investment decision. If future cash flow exceeds a threshold $c^*(n)$ which is determined by the fundamentals, the firm will invest and $u = 1$. If future cash flow is below the threshold $c^*(n)$, the firm will suspend the project and $u = 0$. These two regions are denoted as the "continuation" and "mothball" regions, respectively. The analytical solution of the above Hamilton-Bellman-Jacobi equation can be derived by specifying its function form in both continuation and mothball regions and applying standard boundary conditions at the cash flow threshold. Berk, Green, and Naik (2004) provides detailed derivations, and I reproduce them in Appendix A for reference.

Following standard arguments (i.e., Itô's Lemma), the risk premium of the R&D venture at stage n, $R(n)$, can be derived as follows

$$R(n) = \frac{(\partial V(c,n)/\partial c)c}{V(c,n)}\lambda \quad (1.7)$$

After the project is completed, no further investment decision is needed and the venture is equivalent to a traditional cash producing project which demands the same risk premium λ as the stochastic cash flow process. Before completion, in the mothball region, the firm is equivalent to an option to invest. This is riskier than the underlying cash flow because of the implicit leverage feature of options. In the continuation region, the firm consists of an option to suspend, the discounted value of future cash flow, and the expected R&D cost. Thus the firm in the continuation region is less risky than the mothball region and riskier than the underlying cash flow. In the following subsections, I develop hypotheses that show how a R&D firm's risk premium $R(n)$ varies with its investment level RD and the obsolescence rate ϕ.

1.2.3 R&D investment and risk premium

At any point in time prior to completion, the firm has to make an investment decision. If future cash flow is below a threshold, the firm will suspend the project. A firm that needs to overcome a higher cash flow threshold in order to continue the project is riskier because a high cash flow threshold increases the chance of project suspension in the event of an adverse shock to future cash flow. Therefore, the firm's investment decision and value are more sensitive to the systematic risk the cash flow carries.

A higher R&D investment requirement tends to lower the value of the option to continue the project and thus raises the cash flow threshold the firm needs to overcome. Therefore, a higher R&D investment requirement leads to a higher risk premium. In other words, firms with intensive R&D investments have high expected returns. Furthermore, this positive relation is stronger for firms from industries with high obsolescence rates (i.e., competitive industries) because a high obsolescence rate further raises the cash flow threshold and R&D-intensive firm's investment decisions and value are more sensitive to the systematic risk the cash flow carries. A high obsolescence rate leads to a high probability of cash flow extinguishing and thus lowers the benefit of continuing the R&D investment. Therefore, the obsolescence rate is positively associated with the cash flow threshold.

The above insight yields the first hypothesis.[4]

Hypothesis 1. In the continuation region where cash flow $c >= c^*(n)$,

$$\frac{\partial R(n)}{\partial RD} > 0 \tag{1.8}$$

$$\frac{\partial^2 R(n)}{\partial RD \partial \phi} > 0 \tag{1.9}$$

To have a visual understanding of these effects, I numerically show these relations with a R&D venture which needs five stages to complete. I adopt the same parameter values as in Berk, Green, and Naik (2004), which are as follows: the risk-free rate r is 7% per year. The drift μ and the standard deviation σ of the cash flow process is 3% and 40%, respectively. The risk premium for the cash flow process λ is 8% per year. The success intensity π is set as a constant throughout all stages of the venture for simplicity.[5] The value is 2.0, which is equivalent to 86% probability of successfully completing one stage in a year.[6] The absolesence rate ϕ ranges from 0.05 to 0.25 cross-sectionally.[7] The required R&D investment is set to be a constant throughout all five stages. It is 5.0 for firms with low R&D inputs and 25.0 for firms with high R&D inputs.[8]

[4]These properties are not proved analytically. I show these properties using numerical examples.

[5]The results are robust to how the success intensity varies with each additional completed stage. For example, in one of the unreported cases, I set the value of π in an increasing manner. That is, it starts with a low value and increases by 0.2 as the project advances to the next stage. The results still hold.

[6]If $N(t)$ denotes the number of events that occurred before time t and follows the possion distribution, then the probability that there are k events occurred during the time interval $[t, t+\tau]$ is $P[N(t+\tau) - N(t) = k] = \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}$. Thus $P[N(t+\tau) - N(t) > 0] = 1 - e^{-\lambda\tau}$. Let $\tau = 1$ and $\lambda = 2$, then the probability of completing at least one stage in a year is $1 - e^{-2} = 0.86$.

[7]If $N(t)$ denotes the number of events that occurred before time t and follows the possion distribution, then the probability that there are k events occurred during the time interval $[t, t+\tau]$ is $P[N(t+\tau) - N(t) = k] = \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}$. Thus $P[N(t+\tau) - N(t) = 0] = e^{-\lambda\tau}$. Let $\tau = 1$ and $\lambda = \phi$, then the probability of surviving one year without obolescence is $e^{-\phi}$. The range of ϕ from 0.05 to 0.25 corresponds to a surviving probability range from 0.77 to 0.95.

[8]R&D reporting firms are sorted into quintile portfolios based on their R&D intensites measured by R&D expenditures scaled by assets. The mean value of R&D intensity is 0.01 in the lowest qunintile and 0.24 in the highest quintine. In the model, we can obtain a variable which is equivalent to this R&D intensity by using the required R&D investments scaled by firm value averaged across all stages of the project. When ϕ is 0.1 and the future cash flow is 25, the average ratio of R&D investment to firm value is 0.05 and 0.29 for firms with R&D input equal to 5

Figure 1.1 plots firms' risk premium $R(n)$ against their R&D investment requirements RD for different levels of obsolescence rates (i.e., different levels of competition) and for different stages prior to the completion of the project. The horizontal parts in the plot correspond to the mothball regions where the cash flow is below the threshold. As is shown in the plots, R&D investment is positively related to the risk premium in the continuation regions. This is true for all stages of the R&D venture. Moreover, this positive R&D-return relation is stronger for firms from competitive industries (i.e., industries with higher obsolescence rates). In addition, this positive relation weakens as the project gets closer to completion. As is shown in the bottom right plot when $n = 4$, the magnitude of the risk premium is much smaller than that in earlier stages. This is because firm's value increases as more stages are completed and this reduces the chance of project suspension.

The same intuition is illustrated in Figure 1.2 where the risk premium is computed by averaging it over all stages of the project. All parameter values are the same as in Figure 1.1. As is displayed, the risk premium increases with the level of R&D investment requirement and this positive relation is stronger for firms operating in competitive industries. The indications of this hypothesis is consistent with the empirical findings in the second part of this paper.

1.2.4 Competition and risk premium

A high obsolescence rate leads to a high probability that the potential or actual future cash flow will be extinguished and thus lowers the benefit of continuing the R&D investment. Therefore, firms in industries with higher obsolescence rates (i.e., competitive industries) need to pass higher cash flow thresholds to continue the project and thus have higher risk premium. In other words, firms from competitive industries have higher expected returns. In addition, this positive relation is stronger among firms with intensive R&D inputs which are also positively associated with the cash flow thresholds. This insight yields the second hypothesis.[9]

Hypothesis 2. In the continuation region where cash flow $c >= c^*(n)$,

$$\frac{\partial R(n)}{\partial \phi} > 0 \tag{1.10}$$

$$\frac{\partial^2 R(n)}{\partial \phi \partial RD} > 0 \tag{1.11}$$

Similarly, I use numerical examples to illustrate the properties stated in the above hypothesis. Figure 1.3 plots firms' risk premiums against the obsolescence rates (or levels of competition) for different levels of R&D investment requirements and for different stages. It can be seen clearly that the positive competition-return relation prevails in each subplot of the figure and the relation is more pronounced among firms with intensive R&D investments.

and 25, respectively.

[9] Equation (11) in hypothesis 2 is the same as Equation (9) in Hypothesis 1 in terms of mathematics. I put the same equation in two different hypotheses because there are two relations in the paper: the R&D-return relation and the competition-return relation and I want to illustrate the properties of these two relations separately.

Figure 1.4 shows the same intuition by plotting the risk premium averaged over different stages against the obsolescence rate. As is shown, the positive relation between competition and the risk premium manifests among firms with high R&D inputs. The indications of this hypothesis is also consistent with the empirical findings.

Figure 1.5 and Figure 1.6 illustrate the relations between risk premium and R&D investment and obsolescence rate in integrated three-dimensional plots. Figure 1.5 provides subplots for different stages and Figure 1.6 uses the risk premium averaged across different stages. As is shown, the risk premium increases as R&D investment requirement and obsolescence rate increases. The flat parts in the plot are the mothball regions where cash flow is lower than the threshold. Similar pattern can be found in Figure 1.6.

To check whether the risk premium patterns in the numerical examples still hold with different parameter values, I vary the value of the parameters in the model by 50% and find consistent results. In Figure 1.7 and Figure 1.8, I show the plots for two cases: the volatility of the cash flow process, σ, and the success intensity, π.[10] These are key variables that have large impact on the value of the option to invest and hence firm's investment decisions and demanded risk premium. In the original exercise, σ is set to 40% and π is set to be 2.0. In Figure 1.7, I show the numerical results when σ is set to 0.2 and 0.6, respectively. In Figure 1.8, I present the numerical results when π is set to 1.0 and 3.0, respectively.[11] In both figures I show the relations between the risk premium averaged across all stages and R&D investment requirement (RD) and competition level (obsolescence rate, ϕ) in 3D plots. As is displayed, the plots in Figure 1.7 and Figure 1.8 show similar risk premium patterns as the plot in Figure 1.6.

In sum, the hypotheses that the positive R&D-return relation prevails in competitive industries and that the positive competition-return relation is stronger among R&D-intensive firms follow directly from the Berk, Green, and Naik's (2004) framework for analyzing R&D return dynamics. Next, I turn to test the above hypotheses with a variety of empirical methods.

1.3 Data

1.3.1 Sample Selection and Definition of Variables

My main data sources are from the Center for Research in Security Prices (CRSP) and COMPUSTAT Annual Industrial Files from 1963 to 2009. I obtain firms' monthly stock returns from CRSP and firms' accounting information from COMPUSTAT. To be included in the sample, the firm must have a match in both data sets. Following Fama and French (1992), only NYSE, AMEX-, and NASDAQ-listed securities with share codes 10 and 11 are included in the sample. That is, only firms with ordinary common equity are included (ADRs, REITs, and units of beneficial interest are excluded). Finally, firms in financial and regulated industries are excluded.

[10]The results are robust when I vary the value of other parameters.

[11]The assumption, $\pi = 1$, corresponds to a 63% probability of completing at least one stage in a year. $\pi = 3$, corresponds to a 95% probability of completing at least one stage in a year.

To ensure that the accounting information is already incorporated into stock returns, I follow Fama and French (1992) to match accounting information for all fiscal year ends in calendar year $t-1$ with CRSP stock return data from July of year t to June of year $t+1$. So there is a minimum half a year gap between fiscal year end and the stock return, which provides a certain period of time for the accounting information to be impounded into stock prices. However, firms have different fiscal year ends and thus the time gap between the accounting data and matching stock returns varies across firms. In order to check if this matching procedure biases the main results, I also perform the tests using the sample of firms with December fiscal year ends and similar results are obtained.

Product market competition is measured by Herfindahl index (HHI), a measure commonly used by researchers in the literature of industrial organization.[12] It is defined as the sum of squared market shares

$$HHI_{jt} = \sum_{i=1}^{N_j} s_{ijt}^2 \tag{1.12}$$

where s_{ijt} is the market share of firm i in industry j in year t, N_j is the number of firms in industry j in year t, and HHI_{jt} is the Herfindahl index of industry j in year t. Market share of an individual firm is calculated by using firm's net sales (COMPUSTAT item #12) divided by the total sales value of the whole industry.[13] Following Hou and Robinson (2006), I classify industries with three-digit SIC code from CRSP and all firms with non-missing sales value are included in the sample to calculate the Herfindahl index for a particular industry.[14] The calculation is performed every year and the average values over the past three years is used in the analysis as the Herfindahl index of an industry to prevent potential data errors.[15]

Throughout the paper four measures of R&D intensity are employed to proxy for firm's innovation level: R&D expenditure scaled by total assets (RDA), R&D expenditure scaled by net sales (RDS), R&D expenditure scaled by capital expenditure ($RDCAP$), and R&D capital scaled by total assets ($RDCA$).[16] Following Chan, Lakonishok, and Sougiannis (2001), I compute the R&D capital assuming an annual depreciation rate of 20% over the past five years. Specifically, the equation is as follows:

$$RDC_{it} = RD_{it} + 0.8 * RD_{it-1} + 0.6 * RD_{it-2} + 0.4 * RD_{it-3} + 0.2 * RD_{it-4} \tag{1.13}$$

where RDC_{it} is the R&D capital for firm i in year t and RD_{it-j} is the R&D expenditure j years ago. Although the R&D intensity measures are constructed differently, they are highly correlated with each other. For example, the correlation between RDS and $RDCA$ is 0.73.

[12] See, e.g., Hou and Robinson (2006), Giroud and Mueller (2011). The use of Herfindahl index to measure product market competition is also supported by the theory. See Tirole(1988), pp221-223.

[13] Using the same procedure, the Herfindahl index can be constructed using firms' assets or equity data. These alternative measures produce qualitatively similar test results.

[14] On the one hand, extremely fine industry classification will result in statistically unreliable portfolios. On the other hand, if the classification is not fine enough, firms in different business line will be grouped together.

[15] This averaging procedure is also used in Hou and Robinson (2006).

[16] These measures are commonly used in the literature of innovation. See, e.g., Lev and Sougiannis (1996, 1999), Li (2011), Cohen, Diether, and Malloy (2011).

1.4 Results

Before testing the main hypotheses, I first replicate the empirical results documented in two seminal papers: Hou and Robinson (2006) and Chan, Lakonishok, and Sougiannis (2001). Hou and Robinson (2006) find that firms in competitive industries earn higher abnormal returns than firms in concentrated industries over the sample period of 1963 to 2001. Chan, Lakonishok, and Sougiannis (2001) show that firms conducting intensive R&D activities are associated with higher stock returns over the period of 1975 to 1995. I extend the sample period to 2009 and re-examine these two empirical relations by using portfolio sorting approach.

First I replicate the positive R&D-return relation. In June of each year t, firms are sorted into quintile portfolios based on their value of R&D capital scaled by total assets in year $t-1$.[17] Then time-series regression of excess monthly portfolio return on risk factors is performed using Fama-French three-factor model and Carhart (1997) four-factor model. In particular, the Fama-French three-factor model is estimated in the following equation

$$R_t = \alpha + \beta_1 \times RMRF_t + \beta_2 \times SMB_t + \beta_3 \times HML_t + \epsilon_t \tag{1.14}$$

and the Carhart (1997) four-factor model is estimated as

$$R_t = \alpha + \beta_1 \times RMRF_t + \beta_2 \times SMB_t + \beta_3 \times HML_t + \beta_4 \times UMD_t + \epsilon_t \ , \tag{1.15}$$

where R_t is the excess monthly return of a quintile portfolio, $RMRF_t$ is the value-weighted market return minus the risk free rate in month t, SMB_t is the month t size factor, HML_t is the book-to-market factor in month t, and UMD_t is the month t Carhart momentum factor. $RMRF_t$, SMB_t, and HML_t factors are downloaded from Kenneth French's website and the UMD_t factor is constructed according to Carhart (1997).

Panel A and Panel B in Table 1.1 report the intercept α of the estimation. Both equal-weighted and value-weighted portfolio results are presented. When computing value-weighted return, I use the market capitalization from previous month as the weight. Consistent with Chan, Lakonishok, and Sougiannis (2001) and other studies (e.g., Li (2011), Hirshleifer, Hsu, and Li (2011), Cohen, Diether, and Malloy(2011)), I find the abnormal returns increase monotonically from the lowest R&D intensity portfolio (Quintile 1) to the highest R&D intensity portfolio (Quintile 5). More specifically, in Panel A, for the equal-weighted portfolio, the monthly three-factor α is –0.03% (t–statistic = 0.28) for quintile 1, it is 0.16% (t–statistic = 1.68) for quintile 3, and it is 0.56% (t–statistic = 2.95) for quintile 5. The return spread between quintile 1 and quintile 5 is 59 basis points per month (t–statistic = 3.02). Similarly, in Panel B, using the Carhart (1997) four-factor model, the return spread between the lowest and highest R&D intensity quintile portfolio grows to 70 basis points per month and it is statistically significant (t–statistic = 3.59).

Next I turn to the analysis of the positive competition-return relation. In June of each year t,

[17] This measure is used in Chan, Lakonishok, and Sougiannis (2001), I just follow their choice for the replication.

firms are grouped into quintile portfolios according to their value of Herfindahl index in year $t-1$. I estimate the abnormal return α by running time-series regression of excess monthly portfolio return on common risk factors. Panel C and Panel D report the results for Fama-French three-factor model and Carhart (1997) four-factor model, respectively. As is shown, firms in competitive industries earn higher abnormal returns, while firms in concentrated industries earn lower abnormal returns. The abnormal return decreases almost monotonically as the competition level decreases from quintile 1 to quintile 5.[18] Specifically, in Panel D, the value-weighted four-factor α is 0.06% (t–statistic = 1.91) for the lowest quintile and it decreases to –0.18% (t–statistic = 1.54) for the highest quintile. The abnormal return difference between the highest and lowest quintiles is –0.24% per month and it is statistically significant (t–statistic = 1.91). However, the equal-weighted return spreads are not significant, although the return pattern across the quintiles does show that firms in competitive industries outperform firms in concentrated industries. All in all, the replication results largely confirm that the finding in Hou and Robinson (2006) is still valid with extended sample period.

In sum, the above analysis not only confirms that the two important empirical results documented in the literature survive in an extended sample period, but also indicates that these risk premiums can not be explained by existing common risk factors such as the market factor, the size factor, the value factor and the momentum factor. The model predicts that the positive R&D-return relation strengthens with the level of competition and the positive competition-return relation strengthens with firm's R&D intensity. That is, there is a strong positive interaction effect between R&D intensity and product market competition. In the rest of the paper, I focus on studying this interaction effect from empirical perspectives.

1.4.1 Interaction between R&D Intensity and Product Market Competition

In this section, I investigate the positive interaction effect between production market competition and firm R&D intensity using portfolio analysis and Fama-MacBeth cross-sectional regressions.

Portfolio Analysis

First I study the interaction effect using a conventional double sorting approach. Specifically, in June of each year t, NYSE, Amex, and NASDAQ stocks are divided into three groups using the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of Industry Herfindahl Index in year $t-1$. Meanwhile, independently, firms with non-missing R&D value are grouped into three portfolios based on the breakpoints for the bottom 30%, middle 40%, and top 30% of the ranked values of R&D intensity measures in the previous year.[19] This results in nine portfolios with different characteristics in product market competition and innovation intensity. Monthly equal-weighted and value-weighted returns on the nine portfolios are calculated

[18]Higher Herfindahl index value means lower competition level.

[19]Sorting on those two measures based on quintile breakpoints or performing 3×5 (i.e., 3 on Herfindahl Index, 5 on R&D intensity measure) or 5×3 (i.e., 5 on Herfindahl Index, 3 on R&D intensity measure) sorting produces similar and sometimes stonger results. The results for the 5×5 sorts are presented in Table B.1 in the Appendix. These additional robustness check show that the main results here are not sensitive to the particular way of sorting.

from July of year t to June of year $t+1$, and the portfolios are rebalanced in June each year. Fama-French three-factor model and Carhart (1997) four-factor model are employed to account for the style or risk differences between those portfolios.

The monthly abnormal returns (i.e., the intercept α of the asset pricing model) to the R&D intensity portfolios in competitive (i.e., bottom 30% of the Herfindahl Index) and concentrated (i.e., top 30% of the Herfindahl Index) industries are reported in Table 1.2. Panel A and Panel B in Table 1.2 present the α of the equal-weighted and value-weighted portfolios when R&D intensity is measured by R&D expenditure scaled by net sales (RDS).

As is shown in Panels A and B, the abnormal return to the R&D intensity portfolio increases monotonically with the level of R&D intensity conditioning on that firms are from competitive industries, while the abnormal return to the R&D intensity portfolio is small (often negative) and insignificant when firms are operating in concentrated industries. That is, the positive and statistically significant risk premium associated with R&D intensity only exists for firms operating in competitive industries. This finding holds for both equal-weighted and value-weighted portfolio returns, and for both Fama-French three-factor and Carhart (1997) four-factor model. Observe, for instance, in Panel A, the monthly equal-weighted four-factor α is –0.03% and insignificant (t-statistic = 0.33) in the low R&D intensity group, it is 0.42% and significant (t-statistic = 5.19) in the medium R&D intensity group, and it is 0.63% and significant (t-statistic = 4.08) in the high R&D intensity group. The return spread is 66 basis points per month (t-statistic = 3.92), which translates to 792 basis points annually, while the return spread is negative (–0.40, t-statistic = 0.92) for firms in concentrated industries. Notice that the portfolios are rebalanced every year in the analysis because the variables used to form the portfolios are updated annually, thus the transaction costs associated with the trading strategy is low and the return spread after adjusting for the transaction costs is still economically large.

Chan, Lakonishok, and Sougiannis (2001) propose to use R&D capital which is the weighted average of R&D expenditures over the past five years to proxy for firm innovation intensity. In order to check whether my findings are robust with this measure, I replace R&D expenditure scaled by sales (RDS) with R&D capital scaled by total assets ($RDCA$) and perform the same tests using double sorting approach. The results are presented in Panels C and D in Table 1.2. Similar abnormal return patterns are obtained. That is, the positive R&D-return relation prevails only among firms in competitive industries, while for firms in concentrated industries the R&D premium is insignificantly small or even negative. For instance, in Panel D, the monthly value-weighted four-factor α is 0.09% and insignificant (t-statistic = 0.84) in the low R&D intensity group, it is 0.11% and insignificant (t-statistic = 1.42) in the medium R&D intensity group, and it is 0.55% and significant (t-statistic = 4.36) in the high R&D intensity group. The return spread is 46 basis points per month (t-statistic = 2.50) for firms in competitive industries, while the return spread is negative (–0.18%, t-statistic = 0.34) for firms in concentrated industries.

To make sure that the main findings do not depend on the way the R&D intensity is defined, I also perform the same test with alternative measures proposed by other studies in the literature.

Table 1.3 report the abnormal returns to the double sorted portfolios when I scale R&D expenditure by capital expenditure or total assets to measure R&D intensity ($RDCAP$ or RDA). The results in the table show that changing the deflater does not affect the return pattern at all. Conversely, the return spread across R&D intensity groups in high competition industries becomes larger. Specifically, in Panels A and B, when the deflater is capital expenditure, the monthly equal-weighted return spread for firms in high competition industries is 0.70% (t-statistic = 4.69) and 0.76% (t-statistic = 5.08) for the three-factor and four-factor model, respectively. When the deflater is total assets, in Panels C and D, the monthly equal-weighted return spread for firms in competitive industries is 0.70% (t-statistic = 4.36) and 0.74% (t-statistic = 4.54) for the three-factor and four-factor model, respectively. Note that these numbers are larger than the correponding values when R&D deflated by sales or R&D capital scaled by assets are used to measure R&D intensity in Table 1.2.

In sum, the double sorted portfolio results in Table 1.2 and Table 1.3 provide strong evidence for the hypothesis that the positive R&D-return relation is stronger for firms in competitive industries, while the innovation premium disappears or even becomes negative when firms are operating in concentrated industries. This finding indicates that the positive R&D-return relation identified in earlier studies is the average effect of firms from industries with different competition levels. Furthermore, the fact that R&D-intensive firms in competitive industries earn positive and significant abnormal returns and R&D-intensive firms from concentrated industries earn no or even negative abnormal returns implies that engaging in intensive R&D activities affects firm's risk profiles and thus its returns, but not absolutely. The effect can be significantly different for firms with the same level of R&D intensity, but surrounded by different levels of competition. This paper shows that competition is an important dimension that affects the relation between R&D intensity and stock returns and this finding sheds new light on the understanding of the mechanism behind the positive R&D premium. [20]

Now I turn to test the second half of the hypotheses. That is, the positive competition-return relation is stronger among R&D-intensive firms. And this insight is confirmed by the results reported in Tables 1.4 and 1.5. In June of each year t, firms are independently sorted into three groups (i.e., bottom 30% (Low), middle 40% (Medium), and top 30% (High)) based on their Herfindahl Index and R&D intensity measure in year $t-1$. The monthly equal-weighted and value-weighted abnormal returns adjusting for Fama-French three-factors and Carhart (1997) four-factors are presented in Table 1.4 and Table 1.5, respectively. In each table I show the results for all four R&D intensity measures: RDS, $RDCA$, RDA, $RDCAP$.

As is displayed in the panels, firms in competitive industries (i.e, Low HHI) outperform firms in concentrated industries (i.e., High HHI) by earning higher abnormal returns over the sample period, but only in R&D-intensive group, while the competition premium is small and in most of the cases, negative in low R&D intensity group. The results hold for both equal-weighted and value-weighted portfolios, and for all R&D intensity measures. For example, in Panel A of Table 1.4, in

[20] So far, in the literature researchers have proposed risk-based explanation and investor under-reaction to intangible assets story to explain the positive relation between R&D investments and stock returns.

the high R&D intensity group, the equal-weighted monthly four-factor α is –0.21% and insignificant (t-statistic = 0.54) in low competition industries, it is 0.42% and marginally significant (t-statistic = 1.90) in medium competition industries, and it is 0.63% and highly significant (t-statistic = 4.08) in high competition industries. The return spread is as high as 84 (t-statistic = 2.17) basis points per month, which translates to almost 10% permium annually. While the equal-weighted return spread is –0.22% (t-statistic = 0.67) in low R&D intensity group. Note that the equal-weighted factor-adjusted competition premium in Hou and Robinson (2006) is 36 basis points per month over the period from 1963 to 2001, the return spread identified here is more than twice as large as that in Hou and Robinson (2006). Thus, the results in Table 1.4 and Table 1.5 show that the competition premium only exists among R&D-intensive firms, providing strong supportive evidence for the hypothesis that the competition premium is stronger for firms with intensive R&D investments.

In sum, the findings using this simple double sorting scheme presented in Table 1.2 through Table 1.5 clearly show a positive interaction effect between product market competition and R&D intensity. Specifically, R&D-intensive firms outperform low R&D firms only in competitive industries and firms in competitive industries outperform firms in concentrated industries only among firms with intensive R&D inputs. If the risk associated with R&D activity is priced, the results in this paper not only provide a rational risk-based explanation for the competition premium puzzle, but also tell us that product market competition is an important dimension that can affect the riskiness of R&D investments, which are crucial elements of firm's operating decisions.

In competitive industries, often firms have to enter into innovation races for a new product or technology with many rivals. If any of the competitors successfully completes the development first, the potential cash flows associated with this R&D project will be extinguished for firms that lose the competition and those firms end up with zero or very low returns on the high R&D inputs which are irreversible. The high probability that the cash flows will be extinguished significantly reduces the benefit of investing and thus affects the value and risk of the option to invest. This negative effect is more severe for firms with intensive R&D investments. Therefore, the competition risk or rival risk can have a large impact on the risk profile of R&D-intensive firms. The results presented in this article indicate that this rival risk has important asset pricing implications and potentially drives the positive R&D-return relation.

Fama-MacBeth regressions

In this section, I follow the literature by testing the hypotheses using standard monthly Fama-MacBeth (1973) cross-sectional return regressions with a set of control variables and find supportive evidence for my main hypotheses. In order to identify the effect of R&D intensity on stock returns in different competition groups and the relation between competition and stock returns in different R&D intensity groups, I create dummy variables for firms in low (bottom 30%), medium (middle 40%), and high (top 30%) R&D intensity groups and firms in low (bottom 30%), medium (middle 40%) and high (top 30%) industry competition groups.[21] The variables of interest are the

[21] It is also convinient for the interpretation of the regression coefficients.

interaction terms between the R&D intensity dummy and the competition dummy. Specifically, the equation used to study the competition-return relation in high R&D intensity group is:[22]

$$\begin{aligned} R_{it} &= \alpha_t + \beta_{1t}(RD_{hit} \times HHI_{lit}) + \beta_{2t}(RD_{hit} \times HHI_{mit}) + \beta_{3t}(RD_{hit} \times HHI_{hit}) + \\ &\quad \beta_{4t} \times HHI_{lit} + \beta_{5t} \times HHI_{hit} + \beta_{6t} \times RD_{lit} + \gamma_t Z_{it} + \epsilon_{it} \ , \end{aligned} \tag{1.16}$$

and the equation used to study the R&D-return relation in high competition group is:

$$\begin{aligned} R_{it} &= \alpha_t + \beta_{1t}(HHI_{lit} \times RD_{lit}) + \beta_{2t}(HHI_{lit} \times RD_{mit}) + \beta_{3t}(HHI_{lit} \times RD_{hit}) + \\ &\quad \beta_{4t} \times RD_{lit} + \beta_{5t} \times RD_{hit} + \beta_{6t} \times HHI_{hit} + \gamma_t Z_{it} + \epsilon_{it} \ , \end{aligned} \tag{1.17}$$

where R_{it} is the month t stock return of firm i, RD_l is the dummy variable equal to one if firm i is below the 30th percentile in R&D intensity every year, and zero otherwise. Similarly, RD_m and RD_h are the dummy variables equal to one for stocks in the middle and top R&D intensity groups, respectively. One of these R&D dummy variables are dropped in the estimation because of redundancy. HHI_l is the dummy variable for firms in high competition industries, HHI_m is the dummy variable for firms in medium competition industries, and HHI_h is the dummy variable for firms in low competition industries. One of these competition dummy variables is also dropped from the estimation because of redundancy. Z_{it} is a set of control variables. The element of Z includes book-to-market ratio (i.e., logarithm of book value of equity divided by market capitalization in month $t-1$,[23]), one-month past stock return (to capture the short-term return reversal effect),[24] cumulative return from month $t-12$ to month $t-2$ (to capture the momentum effect),[25] firm size (i.e., logarithm of market capitalization in month $t-1$), and leverage (i.e., long term debt divided by total assets).[26] The above equations are estimated every month over the period from July 1963 to December 2009 and the mean of the monthly estimate of the coefficients for relevant variables are reported.

Table 1.6 summarizes the estimation results. In columns [1] to [4], I study the competition-return relation in the high R&D intensity group using four measures of R&D intensity: RDS, $RDCA$, RDA and $RDCAP$. In these columns, the high R&D intensity dummy is interacted with three (i.e., low, medium, and high) competition dummies. As is shown, the coefficient on the interaction term for the high competition group is positive and significant, while the coefficients on the interaction terms for the medium and low competition groups are small and insignificant or marginally significant. This pattern holds for all four R&D intensity measures. This finding is consistent with my portfolio sorting results that show a significantly positive abnormal return only for the portfolio of high R&D intensity and high competition. Specifically, in column [2], when R&D intensity is

[22] HHI_{mit} dummy, RD_{mit} dummy, and RD_{hit} dummy are not included in this regression specification because that would cause redundancy. If we know the value of $RD_{hit} \times HHI_{lit}$, $RD_{hit} \times HHI_{mit}$ and $RD_{hit} \times HHI_{hit}$, the value of RD_{hit} is also known. If RD_{hit} and RD_{lit} are known, then RD_{mit} is also known. Similarly, RD_{mit}, HHI_{lit}, and HHI_{mit} are not included in Equation (17).

[23] Using market capitalization in other month such as June or December does not affect the estimation results.

[24] See Jegadeesh (1990) for more details.

[25] See Jegadeesh and Titman(1993) for more details.

[26] Using market leverage instead does not affect the results qualitatively.

measured by $RDCA$, the coefficient on the interaction term $R\&D_h * HHI_l$ (i.e., high R&D intensity and high competition) is 0.55% and highly significant (t-statistic = 4.85), while the coefficient on the interaction term $R\&D_h * HHI_m$ (i.e., high R&D intensity and medium competition) is 0.22% and marginally significant (t-statistic = 1.68), and the coefficient on the interaction term $R\&D_h * HHI_h$ (i.e., high R&D intensity and low competition) is –0.03% and insignificant (t-statistic = 0.19). The use of the dummy variables in the estimation makes the results easier to be interpreted: The coefficients on the interaction terms are equivalent to the portfolio abnormal returns in the tables using portfolio sorting approach. Thus the results in column [2] can be interpreted as follows: in competitive industries, when the R&D intensity is measured by $RDCA$, the high R&D intensity portfolio earns an abnormal return of 55 basis points per month on average, similar in magnitude to the α_s in Panel C of Table 1.2. In contrast, in concentrated industries, high R&D intensity portfolio is not associated with significantly positive abnormal returns. This set of results not only show a positive competition-return relation among intensive R&D firms, but also indicate that firms with high R&D inputs are associated with high stock returns only in competitive industries.

In columns [5] to [8], I perform similar estimation, but instead focusing on studying the R&D-return relation in the high competition group. This time, the high competition dummy is interacted with three (i.e., low, medium, and high) R&D intensity dummies. As expected, the coefficient on the interaction term for the high R&D intensity group is positive and significant, while the coefficient on the interaction term for the medium and low R&D intensity groups is smaller or even negative and insignificant or marginally significant. This pattern prevails in each column, which confirms my previous findings using portfolio sorting approach. More specifically, in column [5], when R&D intensity is proxyed by RDS, the coefficient on the interaction term $R\&D_h * HHI_l$ (i.e, high R&D intensity and high competition) is 0.49% and highly significant (t-statistic = 3.46), the coefficient on the interaction term $R\&D_m * HHI_l$ (i.e., medium R&D intensity and high competition) is 0.08% and significant (t-statistic = 1.96), and the coefficient on the interaction term $R\&D_l * HHI_l$ (i.e., low R&D intensity and high competition) is –0.30% and significant (t-statistic = 2.11). This set of results not only show a positive R&D-return relation in competitive industries, but also indicate that competitive industries outperform concentrated industries only among R&D intensive firms.

In Table 1.7, I use a different regression specification to identify the interaction effect between R&D investment and product market competition on stock returns. The estimation equation is:

$$\begin{aligned} R_{it} = {} & \alpha_t + \beta_{1t}(RD_{hit} \times HHI_{lit}) + \beta_{2t}(RD_{hit} \times HHI_{hit}) + \beta_{3t}(RD_{lit} \times HHI_{lit}) + \\ & \beta_{4t}(RD_{lit} \times HHI_{hit}) + \beta_{5t} \times HHI_{lit} + \beta_{6t} \times HHI_{hit} + \beta_{7t} \times RD_{lit} + \\ & \beta_{8t} \times RD_{hit} + \gamma_t Z_{it} + \epsilon_{it} \ , \end{aligned} \tag{1.18}$$

The regression includes four interactions terms, the competition dummies, the R&D intensity dummies[27], and a set of control variables that are the same as those included in Table 1.6. As is

[27]The competition dummies and the R&D intensity dummies are included to control for any direct effect of competition and R&D intensity. Two of them are dropped because of redundancy.

displayed in Table 1.7, a significantly positive coefficient is identified only for the interaction term $RD_{hit} \times HHI_{lit}$. The coefficients for other three interaction terms are smaller in magnitude and insignificant. This confirms the hypothesis that there is a strong interaction effect between R&D intensity and product market competition on stock returns.

I also perform the Fama-MacBeth regression using original variables instead of dummy variables. The estimation equation is as follows:

$$R_{it} = \alpha_t + \beta_{1t}RD + \beta_{2t}HHI + \beta_{3t}RD \times HHI + +\gamma_t Z_{it} + \epsilon_{it} \ , \tag{1.19}$$

The regression includes the proxy for R&D intensity, the measure for competition, the interaction term between them, and a set of control variables that are the same as those included in Table 1.6. The coefficients when R&D intensity is measured by RDS, $RDCA$, RDA and $RDCAP$ are reported in column [1] to column [4], respectively. As is shown, for all four R&D intensity measures, the coefficient on R&D intensity is positive and the coefficient on competition is negative. Moreover, the coefficients on the variable of interest, the interaction term, is negative and significant. Since low HHI value means high competition level, the negative sign on the interaction term indicates positive interaction between competition and R&D intensity. More specifically, in column [1], when R&D intensity is proxied by RDS, the coefficient on the interaction term is –1.15 (t-statistic = 2.35). In column [2], when R&D intensity is proxied by $RDCA$, the coefficient on the interaction term is –1.08 (t-statistic = 2.00).

In sum, the results from this set of Fama-MacBeth cross-sectional return regressions provide strong supportive evidence for my main hypotheses that the R&D-return relation is stronger in competitive industries and the competition-return relation is stronger among R&D-intensive firms. That is, there is a strong positive interaction between R&D intensity and product market competition.

1.4.2 Alternative Asset Pricing Models

In this section, I further estimate the abnormal returns of the double sorted portfolios using alternative asset pricing models. This set of tests serve two purposes. First, this set of tests provide robustness checks for my main findings. Second, it is likely that the abnormal returns identified might be driven by omitted firm characteristics that are correlated with R&D intensity or product market competition but not captured by the employed asset pricing models. Thus the test results may provide information that sheds light on the potential mechanisms driving the return patterns.

Table 1.9 shows the results when I augment Carhart (1997) four-factor model with additional factors proposed by researchers in the asset pricing literature: the liquidity factor of Pastor and Stambaugh (2003),[28] the takeover factor of Cremers, Nair, and John (2009),[29] and the misvaluation factor of Hirshleifer and Jiang (2010).[30] Panel A and Panel B study the R&D-return relation in

[28]The liquidity factor is downloaded from WRDS website.
[29]I am very grateful to Martijn Cremers for sharing the takeover factor data.
[30]This factor is downloaded from the author's website.

low and high competition industries. Panel C and Panel D study the competition-return relation among low and high R&D intensity firms. Here I only tabulate the results when R&D intensity is measured by R&D expenditure scaled by net sales, unreported estimation results when R&D intensity is measured by alternative variables are available upon request.

First, I extend the four-factor model with the liquidity factor of Pastor and Stambaugh (2003), who find that market-wide liquidity is an important asset pricing variable. As is shown in the first row of each panel, adding this additional factor results in economically and statistically similar abnormal return patterns: That is, high R&D intensity firms earn significantly positive abnormal returns only in competitive industries, and firms from competitive industries outperform firms from concentrated industries only among R&D-intensive firms. More specifically, in Panel A, the monthly equal-weighted α for the high R&D intensity portfolio in competitive industries is 0.73% (t-statistic = 4.36), while it is –0.17% (t-statistic = 0.40) for the high R&D intensity portfolio in concentrated industries. In Panel D, the monthly value-weighted α for the high competition portfolio among R&D-intensive firms is 0.41% (t-statistic = 3.09), while it is 0.06% (t-statistic = 0.51) for the high competition portfolio among low R&D firms. Note that those abnormal returns are similar in magnitude to the baseline results using Carhart (1997) four-factor model, indicating that liquidity can not be the reason driving the positive interaction between R&D intensity and product market competition.

Second, I add the takeover factor of Cremers, Nair, and John (2009) who construct this factor based on a target firm's takeover probability. Cremers, Nair, and John (2009) find that the abnormal returns to the Democracy-Dictatorship governance portfolio in GIM become insignificant when they include the takeover factor in the four-factor model. Since product market competition is related to governance,[31] it is likely that the takeover factor can also adjust the abnormal return identified here to some extent. However, surprisingly, adding this additional factor leads to more pronounced patterns. The abnormal return associated with the high R&D intensity-high competition portfolio is significantly higher than the corresponding value when the four-factor model is used. More specifically, for the five-factor model, the equal-weighted and value-weighted monthly α is 0.99% (t-statistic = 4.05) and 0.88% (t-statistic = 4.57), repectively, but the correponding values for the four-factor model is 0.63% (t-statistic = 4.08) and 0.40% (t-statistic = 3.28), respectively. This finding suggests that firms in the high R&D intensity-high competition portfolio are associated with lower takeover probability and load negatively on this factor.

Finally, I test the behavioral explanation for financial market anomalies by adding the misvaluation factor (UMO, undervalued minus overvalued) of Hirshleifer and Jiang (2010) who use firm's debt or equity issuance/repurchases as the signal of investor overvaluation /undervaluation and show that this UMO factor can account for firm's mispricing to some extent. Since R&D projects often are considered to be part of firm's intangible assets and thus difficult to evaluate by outside investors, there have been researchers in the literature proposing that investor undervaluation might be the explanation for the R&D premium. However, adding this factor does not diminish

[31]See, e.g., Giroud and Mueller (2011).

the effects here at all. R&D-intensive firms in competitive industries continue to earn significantly positive abnormal returns. Surprisingly, the magnitude of α is larger. For example, in Panel A, the monthly equal-weighted α is 0.90% (t-statistic = 4.88), 27 basis points higher than the correponding α using the four-factor model. This finding indicates that investor mispricing can not be the complete story behind the positive relation between R&D investments and stock returns.

In sum, the results of this set of tests are consistent with my main findings and those additional factors can not be the potential driver of the strong interaction effect between competition and R&D intensity.

1.4.3 Tests of Alternative Mechanisms

In this section, I continue to test alternative mechanisms that can shed light on the interpretation of my findings by performing subsample studies.

First I explore the idea of financial constraints. Li (2011) document that the positive R&D-return relation only prevails among financially-constrained firms and the positive financial constraint-return relation only exists among R&D-intensive firms. The author argues that financially constrained firms are more likely to suspend the ongoing R&D projects, which makes the venture riskier. This could potentially provide a proper justification to my results if firms in competitive industries are more likely to be financially constrained than firms from concentrated industries. Thus it is necessary to test my results controlling for financial constraints. In order to isolate this potential mechanism, I divide the full sample into financially-constrained and financially-unconstrained subsamples based on the median value of the distribution of KZ index from Kaplan and Zingales (1997) or SA index from Hadlock and Pierce (2010).[32]

Table 1.10 reports the abnormal returns to the double-sorted portfolios in these two subsamples. Panels A and B study the R&D-return relation in competitive and concentrated industries using financially-constrained and financially-unconstrained subsamples created based on KZ index. As is shown, controlling for financial constraints does not weaken my previous results. The abnormal return pattern prevails in both financially-constrained and financially-unconstrained subsamples. That is, R&D-intensive firms are associated with positive and significant abnormal returns only in competitive industries. This indicates that financial constraints can not be the main driver of the positive interaction between industry competition and R&D intensity documented in this paper. In addition, the positive R&D-return relation does not exist in noncompetitive industries for the financially-constrained subsample, indicating that financial constraints don't seem to affect the R&D-return relation once I control for competition. The results in Panels C and D where the subsamples are constructed based on SA index provide the same insights.

Next I turn to the idea of innovation ability proposed by Cohen, Diether, and Malloy (2011) who provide evidence that R&D predicts future returns only when firms have high ability to translate the outcome of those innovation projects into real sales growth. They document a significantly positive

[32]In addition to the KZ index and SA index, I also used the WW index from Whited and Wu (2006) to proxy for financial constraints. Consistent results are obtained.

adjusted return associated with $GOODR\&D$ (i.e., high R&D intensity and high ability) firms. If R&D-intensive firms in competitive industries also have good abilities, then innovation ability could be the story behind the results here. Following Cohen, Diether, and Malloy (2011), I compute the ability as the average of the coefficients from the regressions of sales on past five year R&D investments.[33] Then I create the low (i.e., below 20th percentile) and high (i.e., above 80th percentile) innovation ability subsamples and estimate the abnormal returns to the double-sorted portfolios in each subsample. The results are reported in Table 1.11. Panels A and B study the R&D-return relation in different competition level groups using two different innovation ability subsamples.

Three facts can be inferred from the table. First of all, my main findings still hold in both low and high innovation ability subsamples. That is, the positive and significant abnormal returns are only identified in the portfolio of firms engaging in intensive R&D investments and at the same time in competitive industries. Second, consistent with Cohen, Diether, and Malloy (2011), the positive interaction between R&D intensity and industry competition is stronger in high innovation ability subsample. Specifically, the equal-weighted monthly α of the $R\&D_{high}HHI_{low}$ (i.e., high R&D intensity and high competition) portfolio is 0.76% (t-statistic = 2.15) in the low innovation ability subsample, while it is 1.00% (t-statistic = 2.06) in the high innovation ability subsample, which is 25 basis points higher per month. The return difference is larger in the value-weighted case: the monthly α for the $R\&D_{high}HHI_{low}$ portfolio is 0.83% and significant (t-statistic = 2.19) in the high innovation ability subsample, while it is only 0.44% and insignificant (t-statistic = 1.53) in the low innovation ability subsample. Third, it is interesting to see that the high R&D intensity portfolio from the high-innovation-ability subsample is not associated with positive and significant abnormal return when they are also from noncompetitive industries. This fact indicates that innoation ability does not seem to affect the positive R&D-return relation once I control for competition. In sum, the results of this test provide supportive evidence for my main findings.

1.4.4 Limited Investor Attention

In this section, I turn to examine whether limited investor attention leads to the abnormal return associated with the portfolio of R&D intensive firms in competitive industries. I use firm size, analyst coverage and firm idiosyncratic volatility to proxy for investor attention. Specifically, I split firms into subsamples based on the median value of the distribution of firm size, firm analyst coverage and firm idiosyncratic volatility. If investor attention can explain my main results, I would expect much stronger results among firms with small size, low analyst coverage, and high idiosyncratic volatility. Size is measured by firm market capitalization in June of year $t-1$. Analyst coverage is computed as the monthly average number of unique analysts estimating earnings in year $t-1$. Firm idiosyncratic volatility is computed as the standard deviation of the residuals of the regression of daily excess stock returns on market excess returns over the period of July in year $t-2$ to June in year $t-1$. The equal-weighted monthly abnormal returns to the double-sorted portfolios computed within those subsamples are reported in Table 1.12.

[33] The details of the computation can be found in the description of Table II in Cohen, Diether, and Malloy (2011).

Panels A, B and C study the R&D-return relation in competitive and non-competitive industries. Panels D, E and E study the competition-return relation in high R&D intensity and low R&D intensity firms. According to the results in Table 1.12, limited investor attention can not explain the main findings in my paper since the positive and significant abnormal return to the portfolio of R&D intensive firms in competitive industries exists in both subsamples and the magnitude of the abnormal returns is comparable. For example, when the subsample is constructed based on size, the equal-weighted monthly abnormal return is 0.65% (t-statistic = 2.97) for the small-size subsample and it is 0.87% (t-statistic = 3.54) for the big-size subsample. If the limited attention story is true, we would expect the effect in the small-size subsample to be much stronger than that in the big-size subampl since small size firms are supposed to attract less investor attention. However, the results here indicate that this is not the case. The subsmaple studies using analyst coverage and firm idiosyncratic volatitlity provide similar insights that investor underaction can not justify my findings.

It is highly likely that these abnormal returns are in fact the premium associated with some omitted risk factor related to R&D and competition. To provide further evidence for this risk hypothesis, I continue to investigate whether the portfolio with R&D intensive firms in competitive industries is indeed associated with higher cash flow risk by computing the volatility of the performance measures such as return on assets (ROA), return on equity (ROE) and net profit margin (NPM) of different portfolios. ROA is defined as net income divided by the book value of assets, ROE is defined as net income divided by the book value of common equity, and NPM is defined as net income divided by sales. All performance variables are industry-adjusted, which is calculated by subtracting the industry median from the performance variable. Industry median is computed every year for each of the Fama-French 48 industries. The results are presented in Table 1.13.

The results in the table are consistent with the risk hypothesis. In all cases, the standard deviations of the performance measures are consistently higher for the portfolio of R&D intensive firms in competitive industries. For instance, the volatility of net profit margin for the portfolio of high R&D intensity and high competition is 8.47% points higher than the portfolio of low R&D intensity and low competition. And the difference is significant with a t-statistic of 2.67. Thus the results of this test provide evidence in supportive of the risk hypothesis that the positive and significant abnormal returns to the portfolio of R&D intensive firms in competitive industries are drivern by some unknown risk factors.

1.4.5 Can the R&D Premium Explain the Competition Premium?

The results from previous sections provide strong supportive evidence for the positive interaction effect between competition and R&D intensity by documenting a significant abnormal return for R&D-intensive firms operating in competitive industries. In this section, I turn to test whether the competition premium can be explained by the R&D premium by constructing an additional factor, the innovation factor,[34] and augmenting the existing asset pricing model with this factor to price

[34] Hirshleifer, Hsu, and Li (2011) also construct an innovation factor by using the innovation efficiency measure and document significant positive average returns associated with the factor.

the competition premium documented in Hou and Robinson (2006).

Following the portfolio approach in Fama-French (1993), I construct the innovation factor by sorting firms on size, book-to-market ratio, and the innovation measure ($RDCA$). In June of each year t, NYSE, Amex, and NASDAQ stocks are divided into three size groups using the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of market equity (stock price times shares outstanding from CRSP) in June for NYSE stocks. In each June, I also independently break NYSE, Amex, and NASDAQ stocks into three book-to-market groups based on the breakpoints for the bottom 30%, middle 40%, and top 30% of the ranked values of book-to-market ratio for NYSE stocks. Book-to-market ratio is calculated as the book value of equity in year $t-1$ divided by the market value of equity in December of year $t-1$. Aslo independently, in each June, I sort NYSE, Amex, and NASDAQ stocks into three innovation groups based on the breakpoints for the bottom 30%, middle 40%, and top 30% of the ranked value of $RDCA$ (i.e., R&D capital scaled by total assets).

27 portfolios are formed from the intersections of the three size groups, three book-to-market groups, and three innovation groups. Monthly value-weighted returns on the 27 portfolios are calculated from July of year t to June of year $t+1$, and the portfolios are rebalanced in June of each year. Thus, every month there are nine low $RDCA$ portfolios and nine high $RDCA$ portfolios. The innovation factor, meant to mimic the common variation in stock returns related to firm innovation level, is defined as the difference, each month, between the simple average of the returns on the nine high $RDCA$ portfolios and the simple average of the returns on the nine low $RDCA$ portfolios. I denote this factor by $RDCA$. The triple sort on size, book-to-market ratio, and the R&D intensity measure is to make sure that $RDCA$ is the difference between the returns on high- and low- R&D intensity stock portfolios with similar size and book-to-market characteristics. This way the return difference should be to some extent free of the influence of the size effect and the book-to-market effect, focusing on the return difference of the low and high R&D intensity stocks.

Table 1.14 presents the summary statistics of the constructed innovation factor, Fama-French three factors (i.e., MKT, SML, HML), and Carhart (1997) momentum factor (MOM). Panel A in the table lists some basic statistics of the above five factors. The average monthly return of the innovation factor from July 1963 to December 2009 is 0.37% (t–statistic = 1.70). This confirms the previous results that there is a significant premium associated with R&D investments. Two other facts can be noticed from Panel A as well. First of all, mean of the $RDCA$ factor is higher than that of the size factor and similar to the value factor over my sample period. Second, $RDCA$ factor is almost as volatile as market, size, book-to-market, and momentum factors. Panel B lists the correlation matrix of these factors. The innovation factor is highly positively correlated with the size factor and the correlation is 0.54. It also has a negative correlation of –0.55 with the book-to-market factor. Panel C reports the coefficients of the regressions of the innovation factor on traditional factors using Fama-French three-factor model and Carhart four-factor model. According to the statistics, the intercept α in the three-factor model is 0.47% (t–statistic = 2.89) per month. Controlling for the momentum factors does not reduce the magnitude and significance of the abnormal return α at all. This provides another evidence that the innovaion factor captures cross-sectional variation in stock returns largely independent of the common known risk factors.

Next I use the innovation factor to price the concentration-minus-competition portfolios constructed in the previous section. The time-series return of the concentration-minus-competition portfolios are constructed by subtracting the monthly returns of the competition portfolios from the monthly returns of the concentration portfolios. According to Hou and Robinson (2006) and my replication results in Table 1.1, the concentration-minus-competition portfolio earn a significantly negative abnormal return over both sample periods.

I run time-series regressions of the excess monthly concentration-minus-competition portfolio returns on common risk factors using Carhart (1997) four-factor model and the five factor model which is formed by augmenting the four-factor model with the innovation factor. The regression intercept α and the loadings on the factors are presented in Table 1.15. Panel A and Panel C in the table are the results from the Carhart (1997) four-factor model using value-weighted and equal-weighted concentration-minus-competition portfolio returns, respectively. Panel B and Panel D are the results from the augmented five-factor model using value-weighted and equal-weighted concentration-minus-competition portfolio returns, respectively. According to Panel A, the value-weighted abnormal return to the hedge portfolio is statistically significant after controlling for common risk factors. More specifically, the estimate of α is –0.24% with a t–statistic of 1.91. The estimation results in Panel B reveal that adding the innovation factor to the four-factor model reduces the magnitude of the abnormal return α in Panel A to almost half and brings it to the insignificant level. Specifically, the regression intercept α is –0.14% and insignificant (t–statistic = 1.15) when the innovation factor is constructed using the R&D intensity measure $RDCA$, and it is –0.14% with a t–statistic of 1.22 when the innovation factor is constructed using the R&D intensity measure RDS. Similar insights can be obtained when comparing the equal-weighted results in Panel C and Panel D. As is shown in Panel C, the four-factor alpha is not significant enough (α = –0.13%, and t–statistic = 1.38) and the five-factor model drives its value and the significance level to a much smaller and almost negligible magnitude. For instance, when the innovation factor is constructed using $RDCA$, the five-factor α is –0.05% with a t–statistic of 0.52.

In sum, the innovation factor captures the variation in the return of the concentration-minus-competition portfolio to a large extent, and the test performed in this section provides a rational explanation to the findings documented in Hou and Robinson (2006).

1.5 Conclusion

This article tackles two asset pricing puzzles by testing the joint effect of product market competition and R&D investments on stock returns, providing new perspectives on the positive competition-return relation and R&D-return relation that have drawn a fair amount of attention from economists. In competitive industries, often firms have to enter innovation races for a new product or technology with many rivals. If any of the competitors successfully completes the development first, the potential cash flows associated with this R&D project will be extinguished for firms that lose the competition and those firms end up with zero or very low returns on the high R&D inputs which are irreversible. The high probability that the cash flows will be extinguished significantly reduces the benefit of investing and thus affects the value and risk of firm's option to invest. This negative effect is more severe for firms with intensive R&D investments. Therefore, R&D-intensive firms' risk increases with the level of competition. Conversely, the risk of firms in competitive industries increases with their R&D intensity.

These insights are confirmed by the results from various tests. I find that the positive R&D-return relation only exists among firms from competitive industries, and the competition-return relation only exists among R&D intensive firms. Further tests show that my findings can not be justified by investor limited attention. I provide some evidence in supportive of the risk hypothesis by showing that the portfolio of R&D intensive firms in competitive industries are associated with higher cash flow risk. Furthermore, an innovation factor which is constructed by following the procedure in Fama-French (1993) can explain the competition premium documented in Hou and Robinson (2006) to a large extent. These findings suggest that competition has a significant impact on R&D-intensive firms' risk and return and potentially drives a large portion of the positive R&D-return relation. In addition, these findings shed new light on the understanding of the heretofore puzzling competition premium.

New ideas and new technologies have translated into rising living standards throughout human history and innovation is considered the sheer driving force behind economic growth. More recently more and more technology innovations occur in private sectors. Thus it is important to understand the asset pricing implications of those high R&D investments. This paper brings new perspectives to this literature by proposing competition as an influential factor that affects the risk and return of the R&D ventures.

1.6 References

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyang Zhang, 2006, "The cross-section of volatility and expected returns," *Journal of Finance* 61, 259–299.

Bena, Jan, and Lorenzo Garlappi, 2011, "Strategic Investments, Technological Uncertainty, and Expected Return Externalities," Working paper, University of British Columbia.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, " Optimal investment, growth options, and security returns," *Journal of Finance* 54, 1553–1607.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 2004, "Valuation and return dynamics of new ventures," *Review of Financial Studies* 17, 1–35.

Carhart, Mark M., 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance* 52, 57–82.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, "Corporate in- vestment and asset price dynamics: Implications for the cross-section of returns," *Journal of Finance* 57, 2577–2603.

Carlson, Murray, Engelbert J. Dockner, Adlai Fisher, and Ron Giammarino, 2010, "Leaders, Followers, and Risk Dynamics in Industry Equilibrium," Working paper.

Chambers, Dennis, Ross Jennings, and Robert B. II, 2002, "Excess returns to R&D-intensive firms," *Review of Accounting Studies* 7, 133–158.

Chan, Louis K. C., Josef Lakonishok, and Theodore Sougiannis, 2001, "The Stock market valuation of research and development expenditures," *Journal of Finance* 56, 2431–2456.

Chan K. C., and Nai-Fu Chen, 1991, "Structural and Return Characteristics of Small and Large Firms," *Journal of Finance* XLVI(4).

Chen, Long, Robert Novy-Marx, and Lu Zhang, 2011, "An Alternative Three-Factor Model," Working paper.

Cohen, Lauren, Karl Diether, Malloy Christopher, 2011, " Misvaluing Innovation," Working paper, Harvard Business School.

Cooper, Ilan, 2006, "Asset pricing implications of nonconvex adjustment costs and Ir reversibility," *Journal of Finance* 61, 139–170.

Cooper, Ilan and Richard Priestley, 2011, "Real investment and risk dynamics," *Journal of Financial Economics* 101, 182–205.

Cooper, Michael J., Huseying Gulen, and Michael J. Schill, 2008, "Asset growth and the cross-section of stock returns," *Journal of Finance* 63, 1609–1652.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, "Characteristics, covariances, and average returns: 1929 to 1997," *Journal of Finance* 55, 389–406.

Eberhart, Allan C., William F. Maxwell, and Akhtar R. Siddique, 2004, " An examination of long-term abnormal stock returns and operating performance following R&D increases," *Journal of Finance* 59, 623–650.

Fama, Eugene F., 1990, "Stock Returns, expected returns, and real activity," *Journal of Finance* 45, 1089–1108.

Fama, Eugene F., and Kenneth R. French, 1992, "The Cross-Section of Expected Stock Returns," *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, "Common risk factors in the returns of bonds and stocks," *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1995, "Size and book-to-market factors in earnings and returns," *Journal of Finance* 50, 131–155.

Fama, Eugene F., and Kenneth R. French, 1996, "Multifactor explanation of asset pricing anomalies," *Journal of Finance* 51, 55–84.

Fama, Eugene F., and Kenneth R. French, 2008, "Dissecting anomalies," *Journal of Finance* 63, 1653–1678.

Fama, Eugene F., and James MacBeth, 1973, "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy* 81, 607–636.

Garleanu, Nicolae B., Stavros Panageas, and Jianfeng Yu, 2011, "Technological growth and asset pricing," Working paper.

Garleanu, Nicolae B., Leonid Kogan, and Stavros Panageas, 2009, "The Demographics of Innovation and Asset Returns," working paper, University of California, Berkeley.

Geroski, Paul A., 1990, "Innovation, technological opportunity, and market structure," *Oxford Economic papers* 42, 586–602.

Gomes, Joao F., Leonid Kogan, and Lu Zhang, 2003, " Equilibrium Cross-Section of Returns," *Journal of Political Economy* 111, 693–732.

Hobijn, Bart and Boyan Jovanovic, 2001, " The information-technology revolution and the stock market: Evidence," *American Economic Review* 91, 1203–1220.

Hou, Kewei and David T. Robinson, 2006, "Industry Concentration and Average Stock Returns," *Journal of Finance* 61, 1927–1956.

Hsu, Po-Hsuan, 2009, "Technological innovations and aggregate risk premiums," *Journal of Financial Economics* 94, 264–279.

Kung, Howard and Lukas Schmid, 2011, " Innovation, Growth and Asset Prices," Working paper.

Lev, Baruch, and Theodore Sougiannis, 1996, "The Capitalization, Amortization, and Value-relevance of R&D," *Journal of Accounting and Economics* 21: 107–138.

Lev, Baruch, and Theodore Sougiannis, 1999, "Penetrating the Book-to-market Box: The R&D Effect," *Journal of Business Finance and Accounting* 26: 419–449.

Li, Dongmei, 2011, "Financial Constraints, R&D investment, and Stock Returns," *The Review of Financial Studies* ,September 2011, pg 2974–3007.

Lin, Xiaoji, 2007," Endogenous technological progress and the cross section of stock returns," working paper.

Liu, Xiaolei Laura and Lu Zhang, 2008, "Momentum Profits, Factor Pricing, and Macroeconomic Risk," *The Review of Financial Studies.*

Lyandres, Evgeny, Le Sun, and Lu Zhang, 2008, "The New Issues Puzzle: Testing the Investment-Based Explanation," *The Review of Financial Studies* 21, 2825–2855.

Pastor, Lubos and Pietro Veronesi, 2009, "Technological Revolutions and Stock Prices," *American Economic Review* 99(4): 1451–1483.

Vuolteenaho, Tuomo, "What Drives Firm-Level Stock Returns," *Journal of Finance* LVII(1), 233–263.

Zhang, Lu, 2005, " The Value Premium," *Journal of Finance* 60, 67–104.

1.7 Tables and Figures

Figure 1.1: R&D investment and risk premium

This figure plots the annual risk premium $R(n)$ of a venture that needs five stages to complete against the R&D investment requirement (RD) for firms with different competition levels (obsolescence rate, ϕ) and over different stages. The top left plot is for projects that have completed the first stage ($n = 1$), the top right plot is for projects that have completed the first two stages ($n = 2$), the bottom left plot is for projects that have completed three stages ($n = 3$) and the bottom right plot is for projects that have completed four stages ($n = 4$). The solid and dashed lines correspond to two different levels of obsolescence rates (ϕ): 0.05 for low competition level and 0.25 for high competition level. The parameters in the model are set as follows: the risk-free rate r is 7% per year. The annual drift μ and standard deviation σ of the cash flow process is 3% and 40%, respectively. The risk premium for the cash flow process, λ, is 8% per year. The success intensity π is set as a constant throughout all stages of the venture for simplicity. The value is 2.0, which is equivalent to 86% probability of successfully completing one stage in a year.

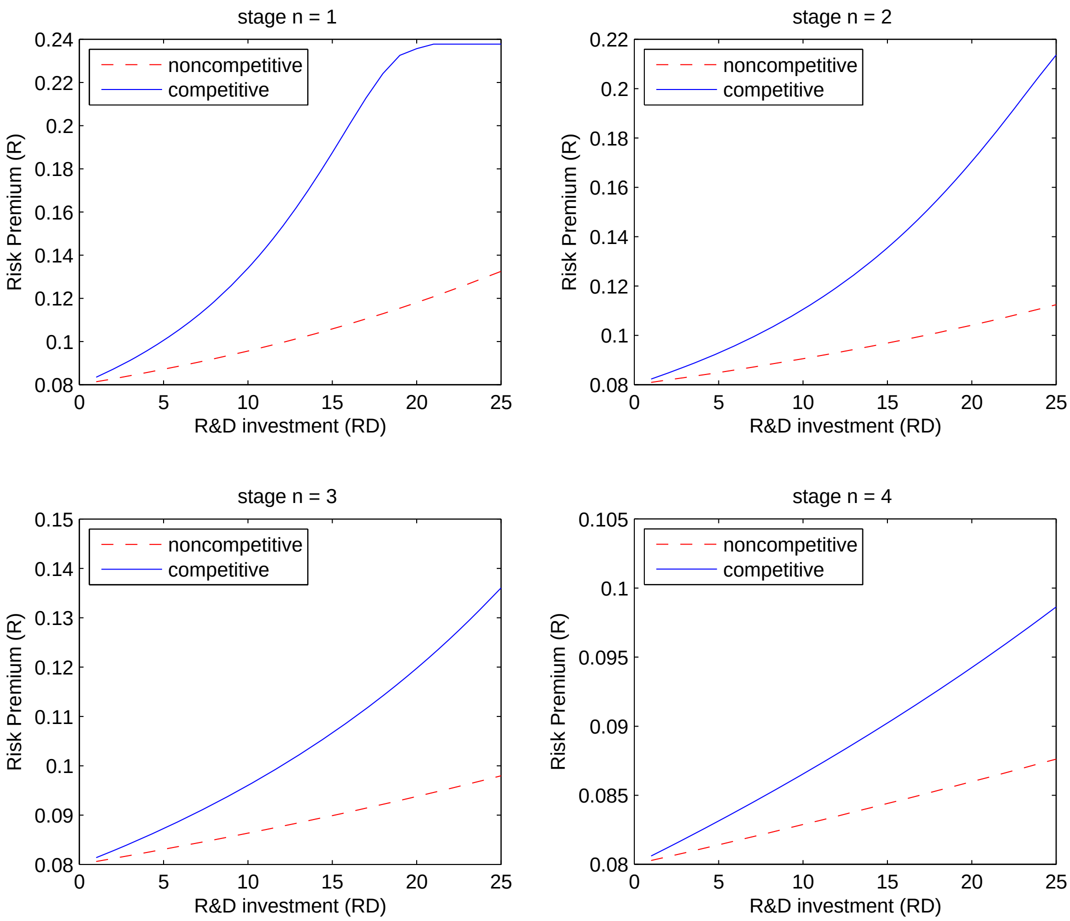


Figure 1.2: R&D investment and averaged risk premium

This figure plots the averaged risk premium $R(n)$ of a venture that needs five stages to complete against the R&D investment requirement (RD) for firms with different competition levels (obsolescence rate ϕ). The dashed and solid lines correspond to two different competition level (obsolescence rate, ϕ): 0.05 for low competition and 0.25 for high competition. The parameters in the model are set as follows: the risk-free rate r is 7% per year. The annual drift μ and standard deviation σ of the cash flow process is 3% and 40%, respectively. The risk premium for the cash flow process, λ, is 8% per year. The success intensity π is set as a constant throughout all stages of the venture for simplicity. The value is 2.0, which is equivalent to 86% probability of successfully completing one stage in a year.

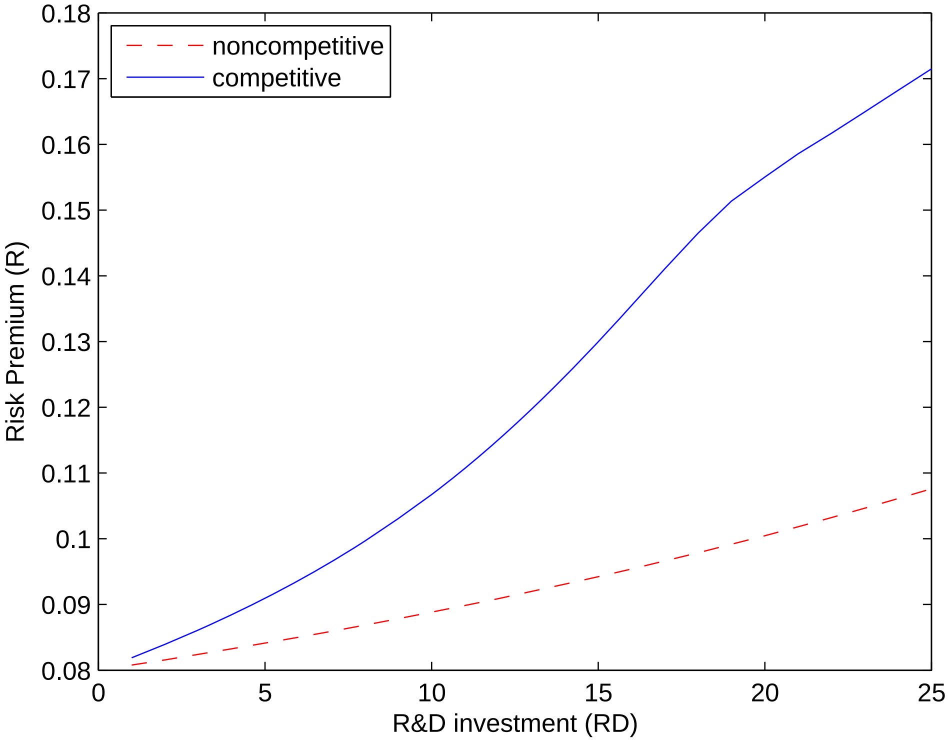


Figure 1.3: Competition and Risk premium

This figure plots the risk premium $R(n)$ of a venture that needs five stages to complete against the competition level (obsolescence rate, ϕ) for firms with different R&D investment requirements (RD). The top left plot is for projects that have completed the first stage ($n = 1$), the top right plot is for projects that have completed two stages ($n = 2$), the bottom left plot is for projects that have finished three stages ($n = 3$) and the bottom right plot is for projects that have completed four stages ($n = 4$). The dashed and solid lines correspond to two different R&D investment level (RD): 5 for low R&D investment and 25 for high R&D investment. The parameters in the model are set as follows: the risk-free rate r is 7% per year. The annual drift μ and standard deviation σ of the cash flow process is 3% and 40%, respectively. The risk premium for the cash flow process ,λ, is 8% per year. The success intensity π is set as a constant throughout all stages of the venture for simplicity. The value is 2.0, which is equivalent to 86% probability of successfully completing one stage in a year.

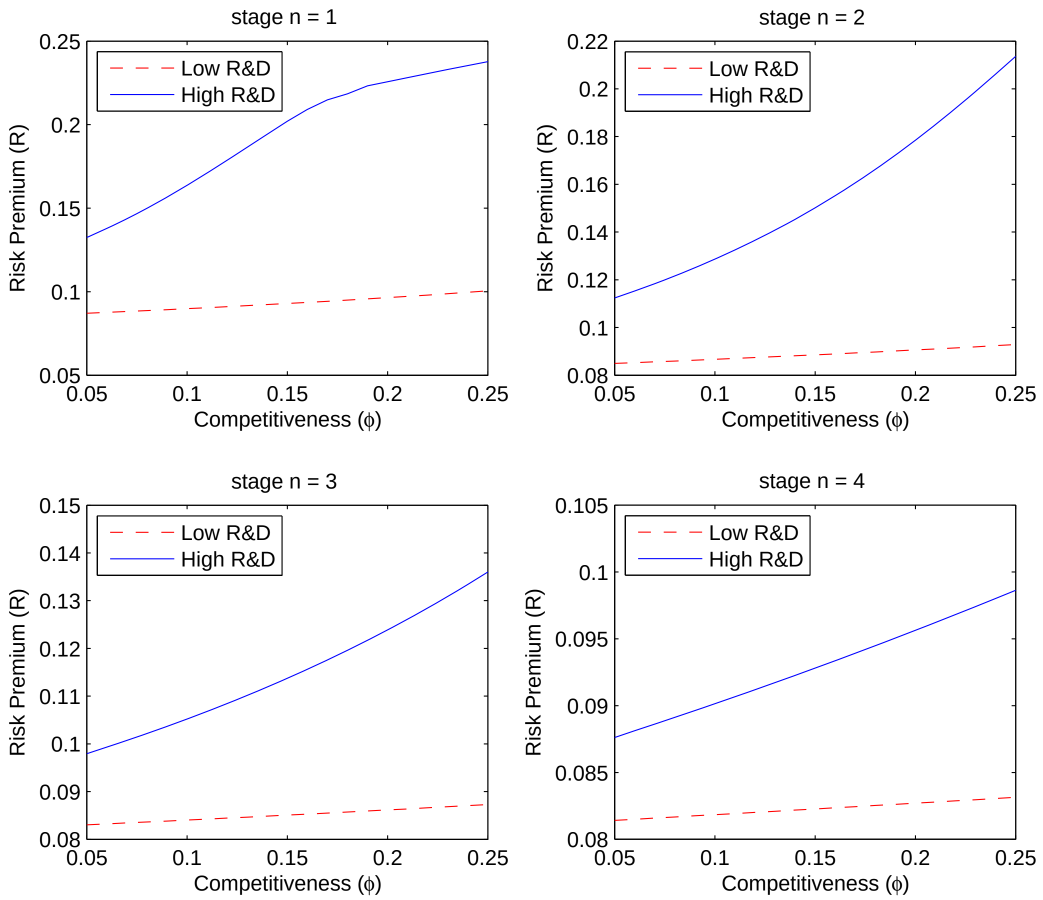


Figure 1.4: Competition and averaged risk premium

This figure plots the averaged risk premium $R(n)$ of a venture that needs five stages to complete against the competition level (obsolescence rate, ϕ) for firms with different R&D investment requirements (RD). The dashed and solid lines correspond to two different R&D investment requirements (RD): 5 for low R&D investment and 25 for high R&D investment. The parameters in the model are set as follows: the risk-free rate r is 7% per year. The annual drift μ and standard deviation σ of the cash flow process is 3% and 40%, respectively. The risk premium for the cash flow process, λ, is 8% per year. The success intensity π is set as a constant throughout all stages of the venture for simplicity. The value is 2.0, which is equivalent to 86% probability of successfully completing one stage in a year.

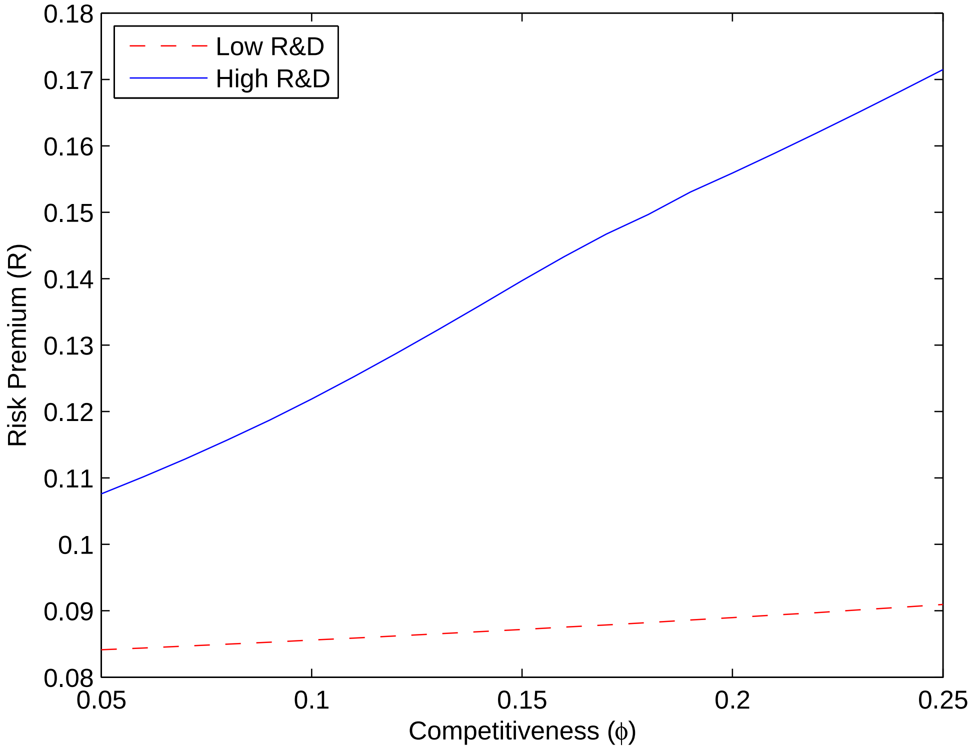


Figure 1.5: Risk premium vs. R&D investment and Competition (3D plot)

This figure plots the risk premium against R&D investment (RD) and competition level (obsolescence rate, ϕ) in 3D for a venture that needs five stages to complete. The top left plot is for projects that have completed the first stage ($n = 1$), the top right plot is for projects that have completed two stages ($n = 2$), the bottom left plot is for projects that have completed three stages ($n = 3$) and the bottom right plot is for projects that have completed four stages ($n = 4$) . The parameters in the model are set as follows: the risk-free rate r is 7% per year. The annual drift μ and standard deviation σ of the cash flow process is 3% and 40%, respectively. The risk premium for the cash flow process ,λ, is 8% per year. The success intensity π is set as a constant throughout all stages of the venture for simplicity. The value is 2.0, which is equivalent to 86% probability of successfully completing one stage in a year.

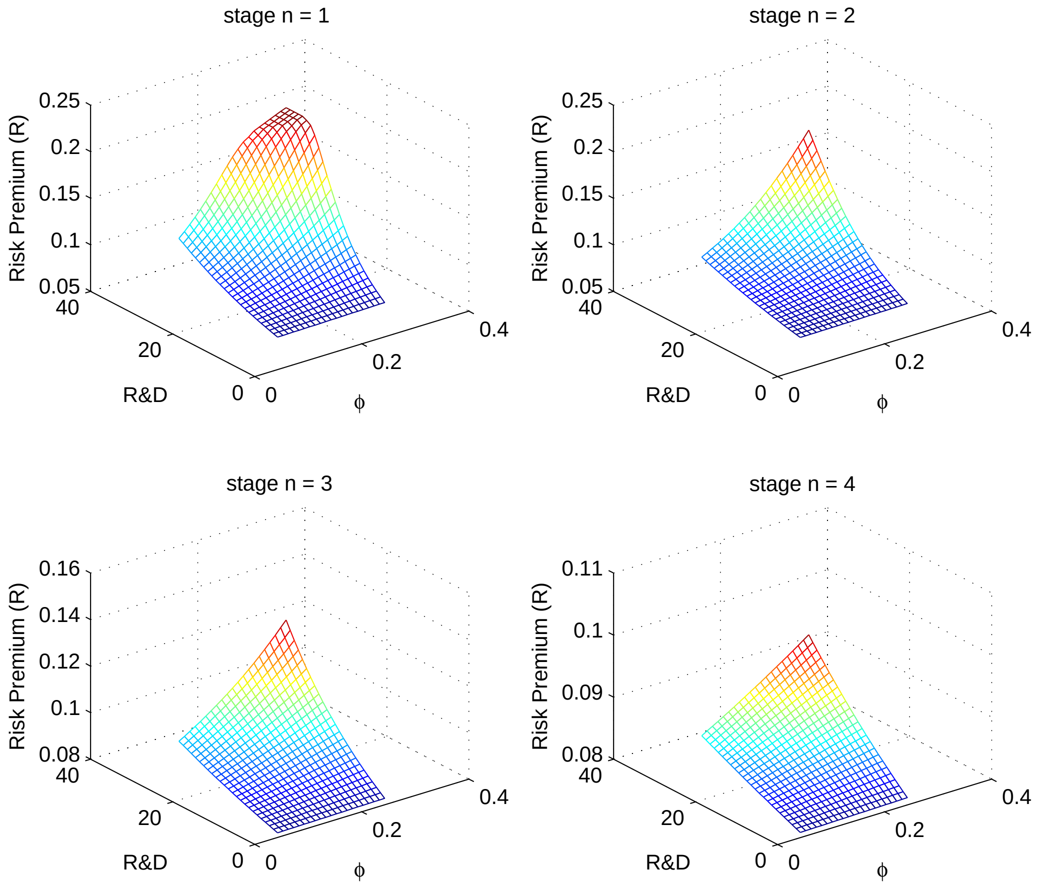


Figure 1.6: Averaged risk premium vs. R&D investment and Competition (3D plot)
This figure plots the risk premium averaged across all stages against R&D investment (RD) and competition level (obsolescence rate, ϕ) in 3D for a venture that needs five stages to complete. The parameters in the model are set as follows: the risk-free rate r is 7% per year. The annual drift μ and standard deviation σ of the cash flow process is 3% and 40%, respectively. The risk premium for the cash flow process ,λ, is 8% per year. The success intensity π is set as a constant throughout all stages of the venture for simplicity. The value is 2.0, which is equivalent to 86% probability of successfully completing one stage in a year.

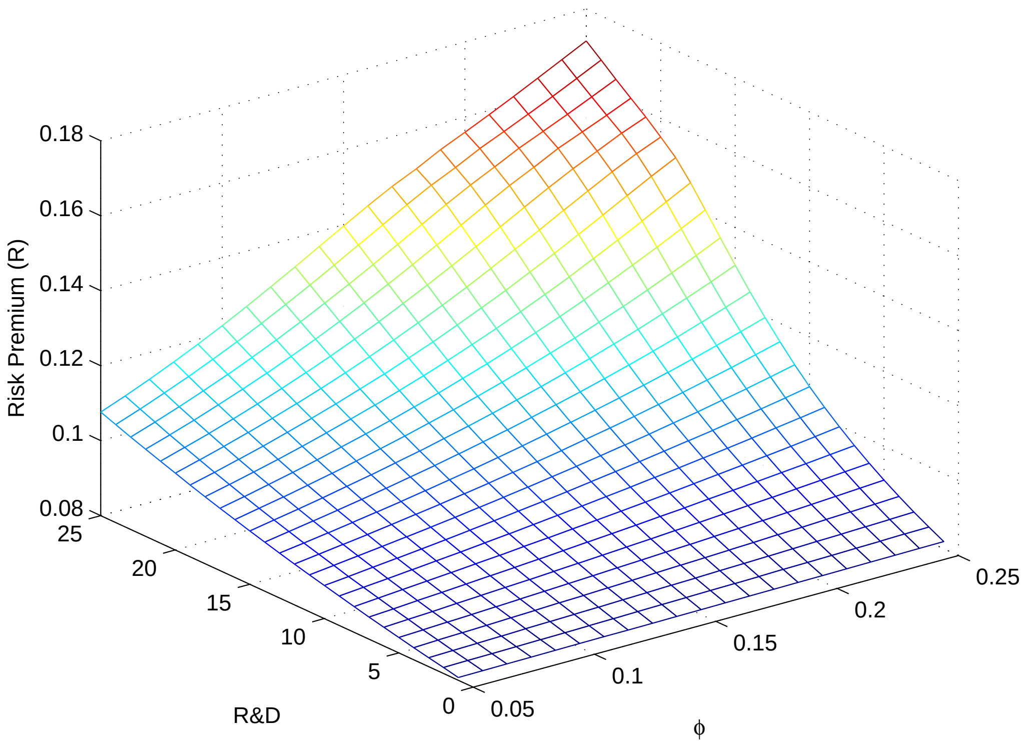
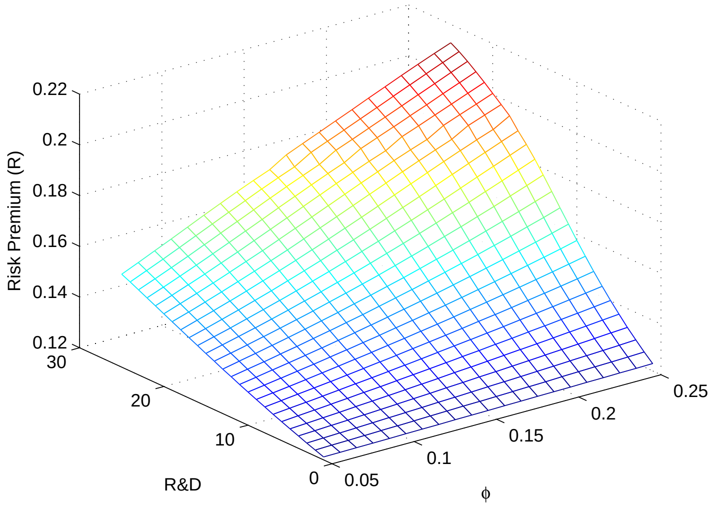


Figure 1.7: 3D plot for averaged risk premium (a. $\sigma = 0.2$ and b. $\sigma = 0.6$)

This figure plots the risk premium averaged across all stages against R&D investment (RD) and competition level (obsolescence rate, ϕ) in 3D for a venture that needs five stages to complete. The standard deviation σ of the cash flow process is set to 20% or 60% instead of 40% in the original case. Other parameters in the model are set as follows: the risk-free rate r is 7% per year. The annual drift μ of the cash flow process is 3% and the success intensity π is set to 2.0 throughout all stages of the venture. (a.) presents the numerical results when $\sigma = 0.2$ and (b.) presents the numerical results when $\sigma = 0.6$.

(a.)



(b.)

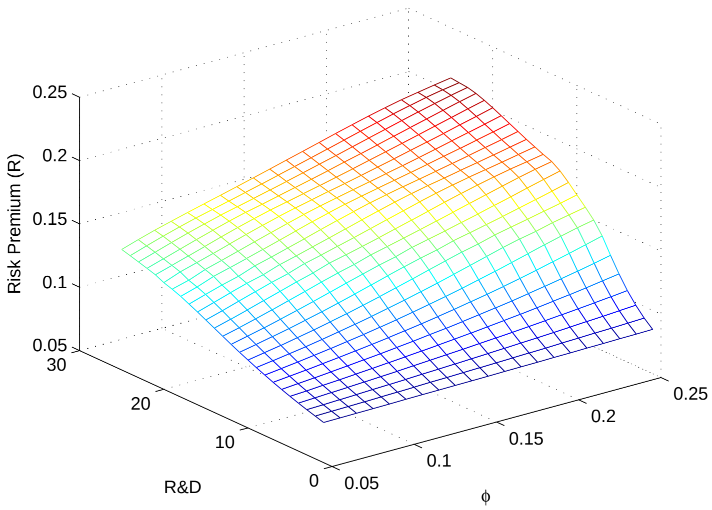
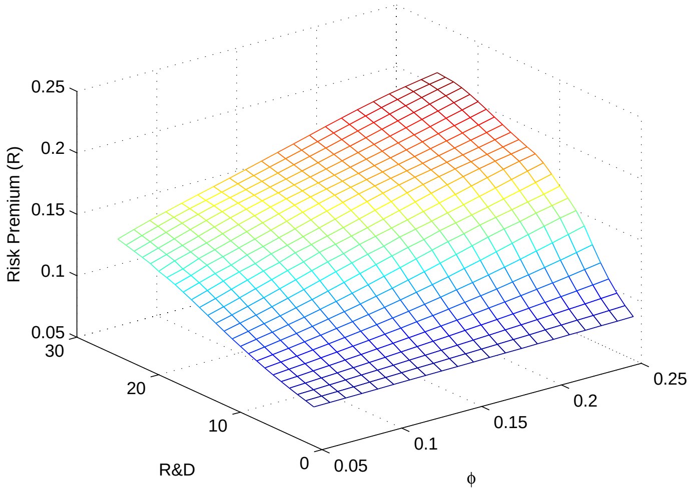


Figure 1.8: 3D plot for averaged risk premium (a. $\pi = 1$ and b. $\pi = 3$)

This figure plots the risk premium averaged across all stages against R&D investment (RD) and competition level (obsolescence rate, ϕ) in 3D for a venture that needs five stages to complete. The success density is set to either 1.0 or 3.0 instead of 2.0 in the original exercise. Other parameters in the model are set as follows: the risk-free rate r is 7% per year. The annual drift μ and the standard deviation σ of the cash flow process is 3% and 40%. respectively. (a.) presents the numerical results when $\pi = 1$ and (b.) presents the numerical results when $\pi = 3$.

(a.)



(b.)

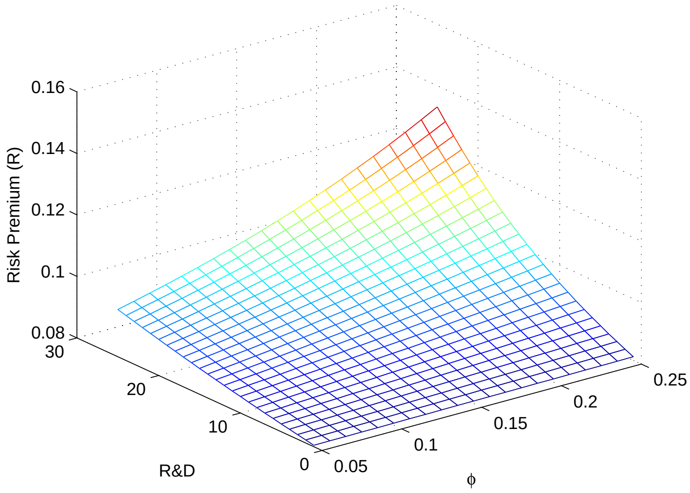


Table 1.1: R&D-Return Relation and Industry Competition-Return Relation

This table reports the monthly abnormal returns (in %) of quintile portfolios sorted on either R&D intensity or product market competition measure. In June of each year t, NYSE, Amex, and NASDAQ stocks are divided into quintile portfolios based on the Industry Herfindahl index in year $t-1$ or R&D intensity measure in year $t-1$. Monthly portfolio abnormal returns are computed by running time-series regression of portfolio excess returns on risk factors in Fama-French three-factor model and Carhart(1997) four-factor model. The measure for product market competition is Herfindahl index computed based on the market share of each firm in the industry. The detailed definition is provided in the data section. The measure for R&D intensity is R&D capital scaled by total assets. Panel A and Panel B present the equal-weighted and value-weihted portfolio returns for the R&D intensity sorted portfolios, respectively. Panel C and Panel D report the results for the Herfindahl index sorted portfolios, respectively. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: R&D-return relation: Equal-Weighted Portfolio Return							
		$R\&D_1(Low)$	2	3	4	$R\&D_5(High)$	$R\&D_5 - R\&D_1$
FF 3-factor	α	–0.03	–0.01	0.16	0.51	0.56	0.59***
	t-stat	(0.28)	(0.17)	(1.68)	(3.87)	(2.95)	(3.02)
Carhart 4-factor	α	0.08	0.07	0.32	0.69	0.79	0.70***
	t-stat	(0.86)	(0.86)	(3.79)	(5.86)	(4.35)	(3.59)
Panel B: R&D-return relation: Value-Weighted Portfolio Return							
		$R\&D_1(Low)$	2	3	4	$R\&D_5(High)$	$R\&D_5 - R\&D_1$
FF 3-factor	α	0.09	–0.09	0.11	0.31	0.50	0.41*
	t-stat	(0.80)	(0.96)	(1.28)	(2.99)	(3.04)	(1.85)
Carhart 4-factor	α	0.08	–0.07	0.18	0.30	0.47	0.40*
	t-stat	(0.68)	(0.70)	(2.13)	(2.82)	(2.85)	(1.76)
Panel C: Competition-return relation: Equal-Weighted Portfolio Return							
		$HHI_1(Low)$	2	3	4	$HHI_5(High)$	$HHI_5 - HHI_1$
FF 3-factor	α	0.01	0.00	–0.06	–0.01	–0.11	–0.12
	t-stat	(0.15)	(0.03)	(0.71)	(0.06)	(0.95)	(1.28)
Carhart 4-factor	α	0.19	0.17	0.06	0.15	0.06	–0.13
	t-stat	(3.29)	(2.38)	(0.69)	(1.70)	(0.56)	(1.38)
Panel D: Competition-return relation: Value-Weighted Portfolio Return							
		$HHI_1(Low)$	2	3	4	$HHI_5(High)$	$HHI_5 - HHI_1$
FF 3-factor	α	0.05	0.03	–0.12	0.02	–0.18	–0.23*
	t-stat	(1.52)	(0.46)	(1.31)	(0.17)	(1.59)	(1.85)
Carhart 4-factor	α	0.06	0.03	–0.15	0.01	–0.18	–0.24*
	t-stat	(1.91)	(0.39)	(1.55)	(0.09)	(1.54)	(1.91)

Table 1.2: Double Sorting on Competition and R&D Intensity: RDS and RDCA

This table reports the monthly abnormal returns (in %) of portfolios sorted on product market competition and R&D intensity. In June of each year t, NYSE, Amex, and NASDAQ stocks are divided into three groups using the breakpoints for the bottom 30%(Low), middle 40%(Medium), and top 30%(High) of the ranked values of Herfindahl index in year $t-1$. Meanwhile, independently, firms with non-missing R&D are grouped into three portfolios based on the breakpoints for the bottom 30%, middle 40%, and top 30% of the ranked values of R&D intensity measure in year $t-1$. Monthly equal-weighted and value-weighted returns on the resulting nine portfolios are calculated from July of year t to June of year $t+1$. Monthly portfolio abnormal returns are computed by running time-series regression of portfolio excess returns on risk factors in Fama-French three-factor model and Carhart(1997) four-factor model. The measure for product market competition is Herfindahl index computed based on the market share of each firm in the industry. The detailed definition is provided in the data section. The measure for R&D intensity is R&D expenditure scaled by net sales (RDS) or R&D capital scaled by total assets ($RDCA$). Panels A and B present the equal-weighted and value-weighted portfolio abnormal returns when the R&D intensity measure is RDS. Panels C and D present the equal-weighted and value-weighted portfolio abnormal returns when the R&D intensity measure is $RDCA$. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
Panel A: Equal-Weighted Portfolio Return : RDS									
FF 3-factor	α	0.01	0.01	–0.49	–0.50	–0.21	0.22**	0.38**	0.59***
	t-stat	(0.04)	(0.06)	(1.23)	(1.15)	(1.97)	(2.27)	(2.24)	(3.50)
Carhart 4-factor	α	0.19	0.13	–0.21	–0.40	–0.03	0.42***	0.63***	0.66***
	t-stat	(1.00)	(0.66)	(0.54)	(0.92)	(0.33)	(5.19)	(4.08)	(3.92)
Panel B: Value-Weighted Portfolio Return : RDS									
FF 3-factor	α	–0.24	0.24	–0.21	0.03	0.00	0.09	0.36***	0.36**
	t-stat	(1.30)	(1.28)	(0.55)	(0.07)	(0.01)	(1.41)	(2.93)	(2.01)
Carhart 4-factor	α	–0.19	0.16	0.00	0.19	0.02	0.12*	0.40***	0.38**
	t-stat	(1.02)	(0.85)	(0.00)	(0.44)	(0.23)	(1.94)	(3.28)	(2.10)

		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
Panel C: Equal-Weighted Portfolio Return : RDCA									
FF 3-factor	α	–0.05	0.20	–0.05	–0.11	–0.03	0.22**	0.56***	0.59***
	t-stat	(0.24)	(1.02)	(0.09)	(0.20)	(0.32)	(2.35)	(3.30)	(3.48)
Carhart 4-factor	α	0.09	0.31	0.22	0.03	0.06	0.38***	0.78***	0.72***
	t-stat	(0.42)	(1.60)	(0.41)	(0.05)	(0.72)	(4.72)	(4.92)	(4.26)
Panel D: Value-Weighted Portfolio Return: RDCA									
FF 3-factor	α	0.15	0.08	–0.04	–0.30	0.08	0.06	0.57***	0.50***
	t-stat	(0.88)	(0.39)	(0.09)	(0.56)	(0.74)	(0.84)	(4.60)	(2.73)
Carhart 4-factor	α	0.16	0.04	0.08	–0.18	0.09	0.11	0.55***	0.46**
	t-stat	(0.92)	(0.21)	(0.16)	(0.34)	(0.84)	(1.42)	(4.36)	(2.50)

Table 1.3: Double Sorting on Competition and R&D Intensity: RDCAP and RDA

This table reports the monthly abnormal returns (in %) of portfolios sorted on product market competition and R&D intensity. In June of each year t, NYSE, Amex, and NASDAQ stocks are divided into three groups using the breakpoints for the bottom 30%(Low), middle 40%(Medium), and top 30%(High) of the ranked values of Herfindahl index in year $t-1$. Meanwhile, independently, firms with non-missing R&D are grouped into three portfolios based on the breakpoints for the bottom 30%, middle 40%, and top 30% of the ranked values of R&D intensity measure in year $t-1$. Monthly equal-weighted and value-weighted returns on the resulting nine portfolios are calculated from July of year t to June of year $t+1$. Monthly portfolio abnormal returns are computed by running time-series regression of portfolio excess returns on common risk factors. The measure for product market competition is Herfindahl index computed based on the market share of each firm in the industry. The measure for R&D intensity is R&D expenditure scaled by capital expenditure or assets ($RDCAP$ or RDA). Panels A and B present the equal-weighted and value-weighted portfolio abnormal returns when the R&D intensity measure is $RDCAP$. Panels C and D present the equal-weighted and value-weighted portfolio abnormal returns when the R&D intensity measure is RDA. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
Panel A: Equal-Weighted Portfolio Return : RDCAP									
FF 3-factor	α	–0.11	–0.05	0.12	0.22	–0.24**	0.16	0.46***	0.70***
	t-stat	(0.65)	(0.21)	(0.29)	(0.49)	(2.37)	(1.56)	(2.74)	(4.69)
Carhart 4-factor	α	0.04	0.05	0.41	0.35	–0.06	0.36***	0.71***	0.76***
	t-stat	(0.21)	(0.22)	(0.99)	(0.77)	(0.68)	(4.36)	(4.65)	(5.08)
Panel B: Value-Weighted Portfolio Return : RDCAP									
FF 3-factor	α	0.08	–0.08	0.01	–0.12	0.04	0.12*	0.48***	0.44**
	t-stat	(0.50)	(0.31)	(0.03)	(0.27)	(0.42)	(1.78)	(3.80)	(2.54)
Carhart 4-factor	α	0.10	–0.16	0.05	–0.11	0.03	0.15**	0.51***	0.48***
	t-stat	(0.62)	(0.58)	(0.12)	(0.24)	(0.31)	(2.34)	(4.05)	(2.78)

		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
Panel C: Equal-Weighted Portfolio Return : RDA									
FF 3-factor	α	–0.11	0.17	–0.58	–0.45	–0.25**	0.18*	0.46***	0.70***
	t-stat	(0.61)	(0.67)	(1.43)	(1.05)	(2.29)	(1.74)	(2.72)	(4.36)
Carhart 4-factor	α	0.05	0.31	–0.32	–0.35	–0.05	0.38***	0.68***	0.74***
	t-stat	(0.29)	(1.21)	(0.80)	(0.81)	(0.58)	(4.70)	(4.44)	(4.54)
Panel D: Value-Weighted Portfolio Return: RDA									
FF 3-factor	α	0.12	0.02	–0.24	–0.39	–0.04	0.09	0.52***	0.56***
	t-stat	(0.77)	(0.09)	(0.60)	(0.89)	(0.41)	(1.25)	(4.38)	(3.21)
Carhart 4-factor	α	0.13	–0.05	–0.14	–0.30	–0.01	0.13*	0.51***	0.52***
	t-stat	(0.83)	(0.18)	(0.35)	(0.69)	(0.12)	(1.84)	(4.28)	(2.98)

Table 1.4: Competition Premium: Equal-Weighted Portfolio

This table reports the monthly equal-weighted abnormal returns (in %) of portfolios sorted on product market competition and R&D intensity. In June of each year t, NYSE, Amex, and NASDAQ stocks are divided into three groups(bottom 30%(Low), middle 40%(Medium), and top 30%(High)) based on its Herfindahl index in year $t-1$. Meanwhile, independently, firms with non-missing R&D are grouped into three portfolios (bottom 30%, middle 40%, and top 30%) according to R&D intensity measure in year $t-1$. Monthly portfolio abnormal returns are computed by running time-series regression of portfolio excess returns on risk factors in Fama-French three-factor model and Carhart (1997) four-factor model. The measure for product market competition is Herfindahl index computed based on the market share of each firm in the industry. Panels A, B, C and D present the results when R&D intensity is measured by *RDS*, *RDCA*, *RDCAP*, *RDCA*, respectively. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: R&D Expenditure scaled by Sales (RDS)

		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
FF 3-factor	α	0.01	–0.11	–0.21*	-0.22	–0.49	0.16	0.38**	0.87**
	t-stat	(0.04)	(0.82)	(1.97)	(1.05)	(1.23)	(0.67)	(2.24)	(2.05)
Carhart 4-factor	α	0.19	0.04	–0.03	–0.22	–0.21	0.42*	0.63***	0.84**
	t-stat	(1.00)	(0.35)	(0.33)	(0.67)	(0.54)	(1.90)	(4.08)	(2.17)

Panel B: R&D Capital scaled by Assets (RDCA)

		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
FF 3-factor	α	–0.05	–0.14	–0.03	0.02	–0.05	0.62***	0.56***	0.61**
	t-stat	(0.24)	(1.05)	(0.32)	(0.46)	(0.09)	(2.73)	(3.30)	(2.01)
Carhart 4-factor	α	0.09	0.00	0.06	–0.03	0.22	0.80***	0.78***	0.57**
	t-stat	(0.42)	(0.04)	(0.72)	(0.32)	(0.41)	(3.64)	(4.92)	(2.12)

Panel C: R&D Expenditure scaled by Capital Expenditure (RDCAP)

		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
FF 3-factor	α	–0.11	–0.18	–0.24**	–0.13	0.12	0.29	0.46***	0.34
	t-stat	(0.65)	(1.44)	(2.37)	(1.49)	(0.29)	(1.33)	(2.74)	(1.60)
Carhart 4-factor	α	0.04	–0.03	–0.06	–0.10	0.41	0.57	0.71***	0.30*
	t-stat	(0.21)	(0.25)	(0.68)	(0.42)	(0.99)	(0.83)	(4.65)	(1.97)

Panel D: R&D Expenditure scaled by Assets (RDA)

		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
FF 3-factor	α	–0.11	–0.21*	–0.25**	–0.14	–0.58	0.43**	0.46***	1.04**
	t-stat	(0.61)	(1.69)	(2.29)	(1.51)	(1.43)	(2.04)	(2.72)	(2.25)
Carhart 4-factor	α	0.05	–0.05	–0.05	–0.10	–0.32	0.62***	0.68***	1.00**
	t-stat	(0.29)	(0.41)	(0.58)	(0.64)	(0.80)	(3.04)	(4.44)	(2.43)

Table 1.5: Competition Premium: Value-Weighted Portfolio

This table reports the monthly value-weighted abnormal returns (in %) of portfolios sorted on product market competition and R&D intensity. In June of each year t, NYSE, Amex, and NASDAQ stocks are divided into three groups (bottom 30%(Low), middle 40%(Medium), and top 30%(High)) based on its Herfindahl index in year $t-1$. Meanwhile, independently, firms with non-missing R&D are grouped into three portfolios (bottom 30%, middle 40%, and top 30%) according to R&D intensity measure in year $t-1$. Monthly portfolio abnormal returns are computed by running time-series regression of portfolio excess returns on risk factors in Fama-French three-factor model and Carhart (1997) four-factor model. The measure for product market competition is Herfindahl index computed based on the market share of each firm in the industry. Panels A, B, C, and D present the results when R&D intensity is measured by *RDS*, *RDCA*, *RDCAP*, *RDCA*, respectively. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
Panel A: R&D Expenditure scaled by Sales									
FF 3-factor	α	–0.24	–0.02	0.00	0.23	–0.21	0.22	0.38***	0.59**
	t-stat	(1.30)	(0.15)	(0.01)	(0.92)	(0.55)	(0.75)	(2.93))	(1.98)
Carhart 4-factor	α	–0.19	–0.06	0.02	0.20	0.00	0.23	0.40***	0.40**
	t-stat	(1.02)	(0.45)	(0.23)	(0.55)	(0.00)	(0.99)	(3.28)	(1.99)
Panel B: R&D Capital scaled by Assets									
FF 3-factor	α	0.15	0.00	0.08	–0.07	–0.04	0.49*	0.57***	0.61**
	t-stat	(0.88)	(0.00)	(0.74)	(0.65)	(0.09)	(1.74)	(4.60)	(2.31)
Carhart 4-factor	α	0.16	–0.04	0.09	–0.07	0.08	0.39	0.55***	0.48**
	t-stat	(0.92)	(0.29)	(0.84)	(0.26)	(0.16)	(1.36)	(4.36)	(2.04)
Panel C: R&D Expenditure scaled by Capital Expenditure									
FF 3-factor	α	0.08	–0.10	0.04	–0.04	0.01	0.18	0.48***	0.48*
	t-stat	(0.50)	(0.78)	(0.42)	(0.71)	(0.03)	(0.68)	(3.80)	(1.75)
Carhart 4-factor	α	0.10	–0.15	0.03	–0.07	0.05	0.26	0.51***	0.46**
	t-stat	(0.62)	(1.09)	(0.31)	(0.50)	(0.12)	(0.97)	(4.05)	(2.01)
Panel D: R&D Expenditure scaled by Assets									
FF 3-factor	α	0.12	–0.04	–0.04	–0.16	–0.24	0.21	0.52***	0.76***
	t-stat	(0.77)	(0.27)	(0.41)	(0.67)	(0.60)	(0.80)	(4.38)	(2.59)
Carhart 4-factor	α	–0.13	–0.08	–0.01	0.12	–0.14	0.20	0.51***	0.65**
	t-stat	(0.83)	(0.55)	(0.12)	(0.35)	(0.35)	(0.76)	(4.28)	(2.07)

Table 1.6: Fama-MacBeth Return Regressions

This table summarizes Fama-MacBeth estimates from monthly cross-sectional regressions of individual stock returns on the R&D dummy, industry competition dummy and the interaction terms between the dummies along with a set of control variables. The estimation equations are given in Section 2.5.3. $R\&D_l$ is the dummy variable equal to one if the firm is below the 30th percentile in R&D intensity every year, and zero otherwise. $R\&D_m$ and $R\&D_h$ are similarly defined. One of these R&D dummy variables are dropped from the estimation because of redundancy. HHI_l, HHI_m and HHI_h are the dummy variables for firms in high, medium, and low competition industries. One of these competition dummy variables is also dropped from the estimation because of redundancy. Control variables include book-to-market ratio (natural log of book value of equity divided by market capitalization in month $t-1$, Ret_{-1}(one-month past return), Ret_{2-12}(cumulative return from month $t-12$ to $t-2$), size (natural log of market capitalization in month $t-1$), and leverage. The mean of the monthly estimate of the coefficients are reported. Columns [1] to [4] study the positive R&D-return relation for firms surrounded by different competition levels. Columns [5] to [8] study competition-return relation in different R&D intensity groups. The variables of interest are the interaction terms between the R&D dummy and industry competition dummy. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	[1] RDS	[2] $RDCA$	[3] $RDCAP$	[4] RDA	[5] RDS	[6] $RDCA$	[7] $RDCAP$	[8] RDA
$R\&D_h * HHI_l$	0.42***	0.55***	0.44***	0.59***	0.49***	0.50***	0.53***	0.57***
	(3.34)	(4.85)	(4.09)	(4.50)	(3.46)	(4.90)	(3.93)	(4.33)
$R\&D_h * HHI_m$	0.13	0.22*	–0.05	0.23				
	(1.15)	(1.68)	(0.24)	(1.58)				
$R\&D_h * HHI_h$	0.21	–0.03	0.16	–0.04				
	(0.92)	(0.19)	(0.74)	(0.18)				
$R\&D_m * HHI_l$					0.08**	0.04	0.05	0.04
					(1.96)	(0.37)	(0.48)	(0.42)
$R\&D_l * HHI_l$					–0.30**	0.06	–0.15*	–0.14
					(2.11)	(0.60)	(1.69)	(1.61)
HHI_l	0.01	0.03	0.01	0.00				
	(0.26)	(0.53)	(0.20)	(0.08)				
HHI_h	–0.05	–0.04	–0.05	–0.03	–0.02	–0.04	–0.02	–0.01
	(0.79)	(0.63)	(0.84)	(0.48)	(0.29)	(0.61)	(0.29)	(0.15)
$R\&D_l$	–0.14**	-0.08	–0.14**	–0.19***	0.07	–0.08	0.02	–0.03
	(2.28)	(1.15)	(2.55)	(3.30)	(0.79)	(0.77)	(0.25)	(0.33)
$R\&D_h$					0.16**	0.21**	0.23***	0.24***
					(2.23)	(2.40)	(2.99)	(3.18)
Size	–0.14***	–0.14***	–0.13***	–0.14***	–0.14***	–0.14***	–0.14***	–0.14***
	(3.24)	(3.23)	(3.12)	(3.15)	(3.27)	(3.31)	(3.20)	(3.18)
ln(B/M)	0.52***	0.51***	0.52***	0.54***	0.53***	0.51***	0.53***	0.54***
	(9.45)	(8.70)	(9.07)	(9.11)	(9.46)	(8.59)	(9.30)	(9.66)
Ret_{-1}	–5.74***	–5.69***	–5.71***	–5.74***	–5.69***	–5.71***	–5.72***	–5.72***
	(14.88)	(14.62)	(14.69)	(14.82)	(14.91)	(14.60)	(14.79)	(14.86)
Ret_{2-12}	0.84***	0.83***	0.83***	0.85***	0.84***	0.83***	0.84***	0.85***
	(6.31)	(6.08)	(6.14)	(6.31)	(6.33)	(6.09)	(6.24)	(6.37)
Leverage	–0.58***	–0.60***	–0.56***	–0.52***	–0.52***	–0.59***	–0.51***	–0.46***
	(3.72)	(3.64)	(3.51)	(3.25)	(3.44)	(3.55)	(3.31)	(3.02)

Table 1.7: Fama-MacBeth Return Regressions (II)

This table summarizes Fama-MacBeth estimates from monthly cross-sectional regressions of individual stock returns on the R&D dummy, industry competition dummy and the interaction terms between the dummies along with a set of control variables. The estimation equations are given in Section 2.5.3. $R\&D_l$ is the dummy variable equal to one if the firm is below the 30th percentile in R&D intensity every year, and zero otherwise. $R\&D_m$ and $R\&D_h$ are similarly defined. One of these R&D dummy variables are dropped from the estimation because of redundancy. HHI_l, HHI_m and HHI_h are the dummy variables for firms in high, medium, and low competition industries. One of these competition dummy variables is also dropped from the estimation because of redundancy. Control variables include book-to-market ratio (natural log of book value of equity divided by market capitalization in month $t-1$, Ret_{-1}(one-month past return), Ret_{2-12}(cumulative return from month $t-12$ to $t-2$), size (natural log of market capitalization in month $t-1$), and leverage. The mean of the monthly estimate of the coefficients are reported. Columns [1] to [4] present the regression results for four different measures of R&D intensity: RDS, $RDCA$, $RDCAP$, RDA. The variables of interest are the interaction terms between the R&D dummy and industry competition dummy. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	[1] RDS	[2] $RDCA$	[3] $RDCAP$	[4] RDA
$R\&D_h * HHI_l$	0.42***	0.50***	0.43***	0.58***
	(3.61)	(3.08)	(3.13)	(3.54)
$R\&D_h * HHI_h$	0.22	0.13	0.25	0.16
	(1.23)	(0.38)	(1.00)	(0.47)
$R\&D_l * HHI_l$	–0.17	0.13	–0.12	–0.09
	(1.55)	(1.06)	(1.15)	(0.87)
$R\&D_l * HHI_h$	0.26	0.35	0.16	0.14
	(1.32)	(1.50)	(0.86)	(0.71)
HHI_l	0.01	0.00	0.02	0.03
	(0.25)	(0.05)	(0.36)	(0.72)
HHI_h	–0.08	–0.08	–0.07	–0.04
	(1.15)	(1.10)	(1.04)	(0.54)
$R\&D_l$	0.05	-0.15	–0.01	–0.07
	(0.45)	(1.30)	(0.06)	(0.72)
$R\&D_h$	0.29***	0.25***	0.27***	0.27***
	(3.33)	(3.17)	(2.58)	(2.74)
Size	–0.14***	–0.14***	–0.14***	–0.14***
	(3.18)	(3.27)	(3.30)	(3.19)
ln(B/M)	0.55***	0.53***	0.51***	0.53***
	(9.72)	(9.50)	(8.66)	(9.33)
Ret_{-1}	–5.73***	–5.75***	–5.71***	–5.72***
	(14.88)	(14.94)	(14.66)	(14.80)
Ret_{2-12}	0.85***	0.84***	0.83***	0.84***
	(6.38)	(6.33)	(6.08)	(6.24)
Leverage	–0.45***	–0.52***	–0.58***	–0.50***
	(2.76)	(3.41)	(3.52)	(3.25)

Table 1.8: Fama-MacBeth Return Regressions (III)

This table summarizes Fama-MacBeth estimates from monthly cross-sectional regressions of individual stock returns on R&D intensity, industry competition and the interaction terms between these two variables along with a set of control variables. The estimation equations are given in Section 2.5.3. Control variables include book-to-market ratio (natural log of book value of equity divided by market capitalization in month $t-1$, Ret_{-1}(one-month past return), Ret_{2-12}(cumulative return from month $t-12$ to $t-2$), size (natural log of market capitalization in month $t-1$), and leverage. The mean of the monthly estimate of the coefficients are reported. Columns [1] to [4] present the regression results for four different measures of R&D intensity: RDS, $RDCA$, $RDCAP$, RDA. The details of the definition of these variables can be found in the data section. The variables of interest are the interaction terms between R&D intensity and industry competition proxy. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	[1] RDS	[2] $RDCA$	[3] $RDCAP$	[4] RDA
$R\&D$	1.68**	1.72**	1.25**	0.92**
	(2.08)	(2.57)	(2.03)	(1.98)
HHI	–0.58*	–0.33	–0.57*	–0.63
	(1.76)	(0.93)	(1.85)	(1.45)
$R\&D * HHI$	–1.15**	–1.08**	–0.93**	–0.76*
	(2.35)	(2.00)	(2.16)	(1.95)
Size	–0.16***	–0.12***	–0.15***	–0.14***
	(3.51)	(2.77)	(3.29)	(3.13)
ln(B/M)	0.48***	0.54***	0.48***	0.57***
	(6.58)	(7.37)	(6.80)	(7.95)
Ret_{-1}	–5.98***	–6.00***	–6.05***	–6.05***
	(13.26)	(11.29)	(13.27)	(13.58)
Ret_{2-12}	0.83***	0.66***	0.78***	0.85***
	(5.12)	(3.86)	(4.65)	(5.34)
Leverage	–0.82***	–0.70**	–0.83***	–0.59**
	(2.75)	(2.19)	(2.84)	(1.99)

Table 1.9: Alternative Asset Pricing Models

This table reports the monthly α (in %) of the portfolios using five-factor asset pricing models. The measure for product market competition is Herfindahl index. The measure for R&D intensity is R&D expenditure scaled by net sales (RDS). The five-factor model is constructed by augmenting the Carhart(1997) four-factor model with additional factors such as the liquidity factor of Pastor and Stambaugh (2003), the takeover factor of Cremers, Nair, and John (2009), and the misvaluation factor of Hirshleifer and Jiang (2010). Panels A and B study the R&D premium in low and high competiton industries. Panels C and D study the competition premium among low and high R&D intensity firms. Due to the data availaility of those additional factors, the sample period is from January 1968 to December 2009 for the liquidity factor, from January 1981 to December 2004 for the takeover factor, and from July 1972 to December 2009 for the misvaluation factor. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
Panel A: Equal-Weighted Portfolio Return									
Liquidity factor	α	0.20	0.23	–0.17	–0.37	0.00	0.45***	0.73***	0.73***
	t-stat	(1.03)	(1.20)	(0.40)	(0.80)	(0.04)	(5.30)	(4.36)	(3.90)
Takeover factor	α	0.11	–0.15	–0.37	–0.48	–0.27**	0.44***	0.99***	1.26***
	t-stat	(0.38)	(0.65)	(0.62)	(0.55)	(1.98)	(3.80)	(4.05)	(3.12)
Misvaluation factor	α	0.07	0.11	–0.58	–0.65	–0.08	0.46***	0.90***	0.98***
	t-stat	(0.32)	(0.58)	(1.29)	(0.96)	(0.79)	(5.04)	(4.88)	(3.69)
Panel B: Value-Weighted Portfolio Return									
Liquidity factor	α	–0.13	0.33*	0.00	0.13	0.06	0.12*	0.41***	0.35**
	t-stat	(0.68)	(1.71)	(0.01)	(0.67)	(0.51)	(1.80)	(3.09)	(2.36)
Takeover factor	α	–0.37	0.13	–0.28	–0.09	–0.04	0.23***	0.88***	0.92***
	t-stat	(1.57)	(0.54)	(0.48)	(1.02)	(0.31)	(2.58)	(4.57)	(3.77)
Misvaluation factor	α	–0.01	0.18	–0.07	–0.06	0.05	0.16**	0.59***	0.54***
	t-stat	(0.04)	(0.99)	(0.16)	(0.23)	(0.41)	(2.15)	(3.99)	(3.02)
		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
Panel C: Equal-Weighted Portfolio Return									
Liquidity factor	α	0.20	0.04	0.00	–0.20	–0.17	0.43*	0.73***	0.90**
	t-stat	(1.03)	(0.34)	(0.04)	(0.59)	(0.40)	(1.91)	(4.36)	(2.35)
Takeover factor	α	0.11	–0.39**	–0.27*	–0.38	–0.37	0.39	0.99***	1.36***
	t-stat	(0.38)	(2.13)	(1.98)	(1.53)	(0.62)	(1.20)	(4.05)	(2.59)
Misvaluation factor	α	0.07	–0.11	–0.08	–0.15	–0.58	0.55**	0.90***	1.48***
	t-stat	(0.32)	(0.82)	(0.79)	(0.41)	(1.29)	(2.23)	(4.88)	(2.63)
Panel D: Value-Weighted Portfolio Return									
Liquidity factor	α	–0.13	–0.04	0.06	0.19	0.00	0.19	0.41***	0.41**
	t-stat	(0.68)	(0.30)	(0.51)	(0.23)	(0.01)	(0.76)	(3.09)	(2.04)
Takeover factor	α	–0.37	–0.28	–0.04	0.33	–0.28	0.49	0.88***	1.16***
	t-stat	(1.27)	(1.28)	(0.31)	(0.59)	(0.48)	(1.23)	(4.57)	(2.60)
Misvaluation factor	α	–0.01	–0.11	0.05	0.06	–0.07	0.43	0.59***	0.66**
	t-stat	(0.04)	(0.68)	(0.41)	(0.34)	(0.16)	(1.55)	(3.99)	(2.15)

Table 1.10: Financial Constraints Subsample Studies

This table reports the monthly α (in %) of the portfolios using subsamples based on financial constraints. The measure for product market competition is Herfindahl index. The measure for R&D intensity is R&D expenditure scaled by net sales (RDS). Panels A and B study the R&D-return relation in low and high competiton industries, respectively. In Panels A and C, firms are restricted to financial-constrainted (above median) firms. In Panels B an D, firms are restricted to financial-unconstrainted (below median) firms. Financial constraint is measured by KZ index from Kaplan and Zingales (1997) or SA index from Hadlock and Pierce (2010). The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
Panel A: Financially-constrained subsample (KZ index)									
Equal-weighted	α	0.11	–0.02	0.41	0.30	0.12	0.40***	0.80***	0.68***
	t-stat	(0.57)	(0.12)	(0.84)	(0.61)	(1.24)	(4.55)	(4.96)	(3.12)
Value-weighted	α	–0.08	–0.03	0.28	0.36	0.15	0.07	0.50***	0.35**
	t-stat	(0.36)	(0.13)	(0.58)	(0.46)	(1.20)	(0.96)	(3.49)	(2.01)
Panel B: Financially-unconstrained subsample (KZ index)									
Equal-weighted	α	0.47	0.31	0.06	–0.41	–0.19	0.37***	0.67***	0.86**
	t-stat	(1.54)	(0.92)	(0.10)	(0.95)	(1.53)	(3.11)	(2.83)	(2.44)
Value-weighted	α	–0.28	0.33	0.24	0.52	–0.22	0.13	0.30	0.52*
	t-stat	(1.13)	(1.20)	(0.41)	(0.67)	(1.63)	(1.05)	(1.52)	(1.82)
Panel C: Financially-constrained subsample (SA index)									
Equal-weighted	α	0.01	0.09	0.61	0.62	0.09	0.36***	0.82***	0.73***
	t-stat	(0.03)	(0.50)	(1.38)	(1.24)	(1.30)	(3.83)	(4.34)	(3.32)
Value-weighted	α	0.00	–0.07	–0.51	–0.55	0.13	0.11	0.52***	0.39*
	t-stat	(0.36)	(0.13)	(0.58)	(0.46)	(0.99)	(1.38)	(3.96)	(1.87)
Panel D: Financially-unconstrained subsample (SA index)									
Equal-weighted	α	0.24	0.23	0.30	–0.02	–0.21*	0.37***	0.77***	0.98***
	t-stat	(0.86)	(0.65)	(0.46)	(0.02)	(1.87)	(3.15)	(3.18)	(4.06)
Value-weighted	α	–0.18	–0.07	–0.31	–0.20	–0.15	0.21	0.34*	0.49*
	t-stat	(0.79)	(0.25)	(0.50)	(0.29)	(0.78)	(1.47)	(1.69)	(1.88)

Table 1.11: Innovation Ability Subsample Studies

This table reports the monthly α (in %) of portfolios using subsamples based on the innovation ability measure proposed in Cohen, Diether, and Malloy (2011). The measure for product market competition is Herfindahl index. The measure for R&D intensity is R&D expenditure scaled by net sales (RDS). Following Cohen, Diether, and Malloy (2011), I compute the ability as the average of the coefficients from the regressions of sales on past five year R&D investments. The details of the computation can be found in the description of Table II in Cohen, Diether, and Malloy (2011). Panels A and B study the R&D-return relation in low and high competiton industries, respectively. In Panel A, firms are restricted to low ability (below 20th percentile) firms. In Panel B, firms are restricted to high ability (above 80th percentile) firms. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
Panel A: Low-innovation-ability subsample									
Equal-weighted	α	–0.28	0.30	–0.24	–0.04	0.25	0.28**	0.76**	0.51**
	t-stat	(0.43)	(0.90)	(0.08)	(0.24)	(0.91)	(2.02)	(2.15)	(2.01)
Value-weighted	α	–0.58	0.67	–0.71	–0.13	–0.04	–0.07	0.44	0.48
	t-stat	(0.89)	(0.12)	(0.24)	(0.55)	(0.15)	(0.47)	(1.53)	(1.67)
Panel B: High-innovation-ability subsample									
Equal-weighted	α	–0.31	0.17	–0.34	–0.03	0.03	0.19	1.00**	0.97**
	t-stat	(0.65)	(0.56)	(1.04)	(0.47)	(0.15)	(1.45)	(2.06)	(2.10)
Value-weighted	α	–0.23	0.33	–0.34	–0.11	0.04	0.08	0.83**	0.79**
	t-stat	(0.47)	(1.10)	(1.04)	(0.53)	(0.19)	(0.50)	(2.19)	(2.00)

Table 1.12: Subsample Studies for Limited Investor Attention

This table reports the monthly α (in %) of the portfolios using subsamples based the median value of the distribution of size, analyst coverage and idiosyncratic volatility. Size is measured by firm market capitalization in June of year $t-1$. Analyst coverage is computed as the monthly average number of unique analysts estimating earnings in year $t-1$. Firm idiosyncratic volatility is computed as the standard deviation of the residuals of the regression of daily excess stock returns on market excess returns over the period of July in year $t-2$ to June in year $t-1$. The measure for R&D intensity is R&D expenditure scaled by net sales (RDS). Panels A, B and C study the R&D-return relation in low and high competiton industries, respectively. Panels D, E, and F study the competition-return relation among low R&D intensity and high R&D intensity firms. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
				Panel A: Size subsample					
Small size	α	0.08	–0.10	–0.08	–0.16	0.07	0.36	0.65***	0.58*
	t-stat	(0.35)	(0.81)	(0.79)	(0.52)	(0.98)	(1.45)	(2.97)	(1.85)
Big size	α	–0.09	0.02	–0.28	–0.17	0.18	0.41*	0.87***	0.69***
	t-stat	(0.73)	(0.21)	(0.65)	(0.33)	(1.11)	(1.75)	(3.54)	(2.58)
				Panel B: Analyst coverage subsample					
Low AC	α	0.20	–0.11	–0.01	–0.21	–0.09	0.27	0.46**	0.55*
	t-stat	(1.10)	(1.21)	(0.57)	(0.83)	(0.78)	(1.58)	(2.23)	(1.83)
High AC	α	0.11	–0.09	–0.25	–0.36	0.08	0.40*	0.84***	0.76***
	t-stat	(0.91)	(0.64)	(1.35)	(1.20)	(0.79)	(1.78)	(3.65)	(2.81)
				Panel C: Idiosyncratic volatility subsample					
High IV	α	–0.05	–0.23	0.03	0.08	–0.01	0.45*	0.71***	0.72**
	t-stat	(0.24)	(0.30)	(0.77)	(0.53)	(0.71)	(1.77)	(3.14)	(2.35)
Low IV	α	0.15	–0.07	–0.15	–0.30	–0.03	0.29	0.69***	0.72**
	t-stat	(0.90)	(0.27)	(0.31)	(0.79)	(0.32)	(1.59)	(2.66)	(2.03)
		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
				Panel D: Size subsample					
Small size	α	0.08	–0.09	0.07	–0.01	–0.08	0.48**	0.65***	0.73**
	t-stat	(0.35)	(0.80)	(0.98)	(0.56)	(0.79)	(2.15)	(2.97)	(1.99)
Big size	α	–0.09	0.03	–0.18	–0.09	–0.28	0.50**	0.87***	1.15**
	t-stat	(0.73)	(0.67)	(1.11)	(0.97)	(0.65)	(2.32)	(3.54)	(2.10)
				Panel E: Analyst coverage subsample					
Low AC	α	0.20	0.16	–0.09	–0.29	–0.01	0.20	0.47**	0.48*
	t-stat	(1.10)	(0.52)	(0.78)	(1.29)	(0.57)	(1.52)	(2.23)	(1.71)
High AC	α	0.11	–0.19	0.08	–0.03	–0.25	0.52*	0.84***	1.09**
	t-stat	(0.91)	(0.43)	(0.79)	(0.70)	(1.35)	(1.97)	(3.65)	(2.21)
				Panel F: Idiosyncratic volatility subsample					
High IV	α	–0.05	–0.23	–0.01	0.04	–0.03	0.40**	0.71***	0.74**
	t-stat	(0.24)	(0.33)	(0.71)	(0.09)	(0.77)	(2.11)	(3.14)	(2.03)
Low IV	α	0.15	–0.09	–0.03	–0.18	–0.15	0.21	0.69***	0.84*
	t-stat	(0.90)	(0.31)	(0.32)	(0.99)	(0.31)	(1.49)	(2.66)	(1.89)

Table 1.13: Cash Flow Risk

This table reports the standard deviation of the profitability measures for four different portfolios: High-R&D-High-Competition, High-R&D-Low-Competition, Low-R&D-High-Competition, and Low-R&D-Low-Competition. High-R&D-High-Competition stands for the portfolio with R&D insensitive firms in competitive industries. High-R&D-Low-Competition stands for the portfolio with R&D insensitive firms in non-competitive industries. Low-R&D-High-Competition stands for the portfolio with Low-R&D-intensity firms in competitive industries. Low-R&D-Low-Competition stands for the portfolio with low-R&D-intensity firms in non-competitive industries. The measure for product market competition is Herfindahl index. The measure for R&D intensity is R&D expenditure scaled by net sales (RDS). The profitability measures are return on assets (ROA), return on equity (ROE), and net profit margin (NPM). Return on assets is defined as net income divided by the value of total assets, return on equity is defined as the ratio of net income to the value of common equity, and net profit margin is defined as net income divided as the value of sales. Those variables are industry-adjusted, which is calculated by subtracting the industry median from the variable. Industry median is computed every year for each of the Fama-French 48 industries. The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

		ROA	ROE	NPM
[1]	High-R&D-High-Competition	8.92	15.90	17.92
[2]	High-R&D-Low-Competition	5.78	12.56	11.78
[3]	Low-R&D-High-Competition	6.71	12.80	12.67
[4]	Low-R&D-Low-Competition	5.17	11.50	9.45
[1]-[2]	Difference	3.14**	3.34**	6.14**
	t-statistic	(2.15)	(2.41)	(2.42)
[1]-[3]	Difference	2.21**	3.10**	5.25**
	t-statistic	(2.06)	(2.20)	(2.40)
[1]-[4]	Difference	3.75***	4.40**	8.47***
	t-statistic	(3.39)	(2.55)	(2.67)

Table 1.14: Summary Statistics of Factors

This table lists summary statistics of the innovation factor, Fama-French three factors (from Kenneth French's website), and Carhart momentum factor (constructed according to Carhart (1997). The five factors are denoted by $RDCA$ (innovation factor), MKT (market factor), SML (size factor), HML (value factor), and MOM (momentum factor), respectively. In June of each year t, NYSE, Amex, and NASDAQ stocks are divided into three size groups using the breakpoints for the bottom 30%(Low), middle 40%(Medium), and top 30%(High) of the ranked values of market equity (price times shares outstanding from CRSP) in June for NYSE stocks. In each June, I also independently break NYSE, Amex, and NASDAQ stocks into three book-to-market groups based on the breakpoints for the bottom 30%, middle 40%, and top 30% of the ranked values of book-to-market ratio for NYSE stocks. Book-to-market ratio is calculated as the book value of equity in year $t-1$ divided by the market value of equity in December of year $t-1$. Aslo independently, in each June, I sort NYSE, Amex, and NASDAQ stocks into three innovation groups based on the breakpoints for the bottom 30%, middle 40%, and top 30% of the ranked value of $RDCA$, which is defined as R&D capital scale by total assets. 27 portfolios are formed from the intersections of the three size groups, three book-to-market groups, and three innovation groups. Monthly value-weighted returns on the 27 portfolios are calculated from July of year t to June of year $t+1$, and the portfolios are rebalanced in June of each year. Thus, every month there are nine low *R&D* portfolios and nine high *R&D* portfolios. The innovation factor is defined as the difference, each month, between the simple average of the returns on the nine high *R&D* portfolios and the simple average of the returns on the nine low *R&D* portfolios. Panel A lists some basic statistics of the five factors. $SKEW$ and $KURT$ refer to skewness and kurtosis, respectively. Panel B lists the correlation matrix of these factors. Panel C reports the coefficients of the regressions of innovation factor on traditional factors using two different asset pricing models: Fama-French three-factor model, and Carhart (1997) four-factor model. The sample period is from July 1963 to December 2009.

Panel A: Basic descriptive statistics

	Mean	T–stat	STD	SKEW	KURT
MKT	0.45	2.36	4.39	-0.50	2.11
SMB	0.25	1.79	3.24	0.52	5.56
HML	0.42	3.26	2.91	0.07	2.49
MOM	0.49	2.16	5.19	-2.64	21.93
RDCA	0.37	1.70	4.94	2.20	19.57

Panel B: Correlation matrix of factors

	MKT	SMB	HML	UMD	RDCA
MKT	1.00				
SMB	0.30	1.00			
HML	-0.41	-0.27	1.00		
MOM	-0.15	-0.16	-0.02	1.00	
RDCA	0.36	0.54	-0.55	-0.01	1.00

Panel C: Regression coefficients of innovation factor on other factors

	α	MKT	SMB	HML	UMD
FF three-factor	0.47	0.07	0.62	-0.70	
	(2.89)	(1.78)	(11.93)	(11.62)	
Carhart four-factor	0.55	0.05	0.59	-0.73	-0.12
	(3.42)	(1.24)	(11.44)	(12.11)	(3.76)

Table 1.15: Pricing the concentration-minus-competition portfolios

This table reports the intercept α from the time-series regressions of monthly concentration-minus-competition portfolio returns on risk factors using Carhart (1997) four-factor model and the five factor model which is formed by augmenting the four-factor model with the innovation factor. I construct quintile portfolios based on the Herfindahl index of each industry which is classified by the three-digit SIC codes and calculate both equally-weighted and value-weighted portfolio returns. The time-series return of the concentration-minus-competition portfolios are constructed by subtracting the monthly returns of the competition portfolios from the monthly returns of the concentration portfolios. $RDCA$ and RDS stand for the innovation factor constructed using R&D capital scaled by total assets and R&D expenditure scaled by sales, respectively. Panel A and Panel C are the results from the four-factor model using value-weighted and equal-weighted concentration-minus-competition portfolio returns, respectively. Panel B and Panel D are the estimation results using the augmented five-factor model. The data of the risk factors, risk-free rate are downloaded from Kenneth-French's website. The momentum factor are constructed according to the procedure in Carhart (1997). The sample period is from July 1963 to December 2009.

Panel A : four-factor model (Value-Weighted Return)							
	α	Mkt_RF	SMB	HML	UMD		
estimate	–0.24*	0.05	0.12	0.41	0.01		
t-stat	(1.91)	(1.60)	(2.96)	(8.99)	(0.54)		
Panel B : five-factor model (Value-Weighted Return)							
	α	Mkt_RF	SMB	HML	UMD	RDCA	RDS
estimate	–0.14	0.06	0.22	0.29	–0.01	–17.53	
t-stat	(1.15)	(1.95)	(5.12)	(5.66)	(0.31)	(5.37)	
estimate	–0.14	0.07	0.22	0.22	–0.02		–21.26
t-stat	(1.22)	(2.50)	(5.45)	(4.35)	(0.73)		(6.97)
Panel C : four-factor model (Equal-Weighted Return)							
	α	Mkt_RF	SMB	HML	UMD		
estimate	–0.13	–0.04	–0.06	0.10	0.02		
t-stat	(1.38)	(1.66)	(1.79)	(2.93)	(0.85)		
Panel D : five-factor model (Equal-Weighted Return)							
	α	Mkt_RF	SMB	HML	UMD	RDCA	RDS
estimate	–0.05	–0.03	0.04	–0.01	0.00	–15.40	
t-stat	(0.52)	(1.36)	(1.07)	(0.15)	(0.10)	(6.06)	
estimate	–0.06	–0.02	0.03	–0.05	–0.01		–17.23
t-stat	(0.64)	(0.85)	(0.94)	(1.18)	(0.46)		(7.24)

Chapter 2

Governance and Equity Prices: Does Transparency Matter?

2.1 Introduction

Since Jensen and Meckling (1976), economists have devoted much effort to studying a firm's governance, which balances the allocation of power between managers and shareholders, and a firm's information environment, which provides shareholders with the data necessary to assess their firm's performance. Clearly, corporate governance and accounting transparency are not only important for academics and managers, but also for regulators. Recent cases of poor governance as, e.g., in the scandals of Enron or Worldcom, lead legislators to mandate new rules that enforce more transparency, suggesting governance and transparency are regarded as substitutes. Accordingly, the need to provide managers with incentives through governance—and thus the benefits of governance—should be smaller for more transparent firms.

Gompers, Ishii, and Metrick (2003, GIM) propose takeover vulnerability as a one specific governance measure and construct the G-index, which consists of 24 anti-takeover and shareholder rights provisions. In their seminal article, GIM uncover an important link between governance and firm performance, since a long-short portfolio that buys good governance firms ("Democracy firms") and sells weak governance firms ("Dictatorship firms") earns a monthly abnormal return of 0.71%. Based on these authors' work, a rich literature has emerged to examine various aspects that explain or even strengthen the effect of governance mechanisms on firm performance.[1]

Apart from the aforementioned anecdotal evidence, we know very little about how governance and transparency are related to firm performance. This paper focuses on a more specific question: Are firms with better governance (measured by a lower G-index) associated with better performance, on average, if they are also more transparent (e.g., measured by forecast dispersion)?[2]

[1] See, e.g., Cremers and Nair (2005), Core, Guay, and Rusticus (2006), Ferreira and Laux (2007), Cremers, Nair, and Peyer (2008), Bebchuk, Cohen, and Ferrell (2009), Cremers, John, and Nair (2009), and Giroud and Mueller (2011).

[2] On the one hand, there is evidence on a positive association between governance and transparency at the international level (see, e.g., Bushman, Piotroski, and Smith (2004), Leuz and Oberholzer (2006), Doidge, Karolyi, Stulz (2007), and Lang, Lins, and Maffett (2009)). On the other hand, Armstrong, Balakrishnan and Cohen (2012) document for U.S. firms affected by state anti-takeover laws between 1985 and 1991 that their information

To answer this question, we study the joint effect of a firm's information environment (or transparency) and its governance on equity returns. We use the G-index developed by GIM to proxy for a firm's governance and measure a firm's information environment by three transparency proxies: forecast error, forecast dispersion, and revision volatility.[3] Every year, we divide GIM's "Democracy firms" (with strong corporate governance) and "Dictatorship firms" (with weak corporate governance) into three equal-sized portfolios based on whether the value of an analyst variable is in the lowest, medium, and highest tercile of its empirical distribution. The GIM-based trading strategy (or hedge portfolio) buys good governance firms and sells bad governance firms in each of the three terciles. To compare the success of the trading strategy across transparency terciles, we follow standard practice and compute returns of the "Democracy-Dictatorship" hedge portfolio adjusted for the Carhart (1997) four-factor model within each of the transparency terciles.

Our main finding is that governance and transparency reinforce each other in that transparent firms benefit more from good corporate governance than opaque firms. Specifically, we find that the hedge portfolio in the high transparency tercile earns a high and significant abnormal return over the sample period, the hedge portfolio in the medium tercile earns a small and insignificant (or marginally significant) abnormal return, and the hedge portfolio in the low transparency tercile earns a smaller and insignificant abnormal return. This pattern applies to each of the three measures of transparency and survives using various deflators of the transparency proxies (i.e., share price, book assets per share, and absolute value of forecast mean). Moreover, when we combine all the information contained in forecast error, forecast dispersion, and revision volatility by constructing their first principal component and use the computed factor as a proxy for a firm's information environment, the hedge portfolio focusing on transparent firms earns a monthly alpha of 1.37% with t-statistic of 3.52 for value-weighted (1.28% with t-statistic of 3.50 for equal-weighted) portfolios, which is nearly twice as large as the abnormal return on governance reported by GIM, while the monthly abnormal return of the hedge portfolio focusing on opaque firms is 0.04% with t-statistic of 0.09 for value-weighted (0.04% with t-statistic of 0.11 for equal-weighted) portfolios.[4] We find similarly strong results when using each of the three transparency proxies individually and when scaling them by lagged assets per share, lagged share price, or absolute value of forecast mean. Our result also holds even if we employ *time-invariant* sample averages of the transparency proxies, which are largely outside of managers' discretion and hence more *permanent* firm characteristics, to construct transparency terciles only once instead of rebalancing them annually. The complementary effect between transparency and governance is also confirmed by profitability measures such as return on assets, return on equity, and net profit margin. We find that the positive relation between good

environment improves when protection from hostile takeovers increases, but they do not study firm performance. Yet, there is *no* reliable relation between governance and transparency in our sample of U.S. firms (see Section 2.3.2), which is consistent with other studies (see, e.g., Larcker, Richardson, and Tuna (2007)). Finally, Hermalin and Weisbach (2007) argue the relation between governance and transparency is more subtle than previously believed. See Section 2 for theoretical arguments and testable hypotheses.

[3]These are all standard measures used frequently by researchers in finance and accounting; see, e.g., Givoly and Lakonishok (1979), Lang and Lundholm (1996), Thomas (2002), and Zhang (2006). Importantly, using accruals quality to measure transparency instead reinforces the interpretation of our findings, as the economic effects are stronger.

[4]This finding for abnormal returns is not inconsistent with the fact that more transparent firms have, on average, a lower cost of equity capital (see, e.g., Leuz and Verrechia (2000), Leuz and Hail (2006), and Botosan and Plumlee (2002)). For instance, the monthly equal-weighted portfolio return is, on average, 1.25% for the lowest, 1.35% for the medium, and 1.54% for the highest tercile of forecast dispersion in our sample during the 1990–2006 period.

corporate governance and operating performance is significant only when firms are transparent.[5]

Consistent with our main result, we find support for the view that transparency facilitates takeovers by comparing average transparency levels of target firms and all firms in the sample over the 1990–2006 period and by estimating an empirical logit model for takeover probability. In the latter case, only good governance firms, which are transparent, are more likely takeover targets.[6]

Our main finding survives numerous robustness tests. We replace the three transparency proxies by accruals quality and find that the economic magnitude of the Democracy-Dictatorship portfolio's abnormal returns for transparent firms increases. Next, we replace G-index by an alternative governance measure (i.e., the E-index), exclude "new economy" firms from the sample, extend the sample period to 2011. In all cases, the results are consistent with our main finding. We also experiment with dividing the full sample into high-institutional-ownership and low-institutional-ownership subsamples or into high-industry-competition and low-industry-competition subsamples. These additional tests not only support our hypothesis, but are also consistent with the findings in Cremers and Nair (2005) and Giroud and Mueller (2011).[7] Moreover, we re-estimate alphas when splitting the sample at median asset size and median leverage ratio to verify that the main result is not largely due to these firm characteristics. Given that smaller firms and firms with less leverage are relatively easier to acquire, our main finding's interpretation is also confirmed because the Democracy-Dictatorship portfolio earns higher abnormal returns in the small firm and the low leverage subsamples. If a large part of the Democracy portfolio is comprised of firms from high abnormal return industries and a large part of the Dictatorship portfolio is formed by firms from low abnormal return industries, then this could blur our identification. We dissolve this concern by using industry-adjusted returns with different asset pricing models. In addition, we provide an integrated test to establish the unique ability of transparency to influence abnormal returns of the Democracy-Dictatorship portfolio using multivariate regression analysis, which allows us to control simultaneously for various variables used in prior research, such as competition, institutional ownership, etc. Finally, we experiment with alternative asset pricing models that include, e.g., the liquidity factor of Pastor and Stambaugh (2003) or the takeover factor of Cremers, John, and Nair (2009) as a fifth factor. Therefore, we conclude that a firm's information environment is an independently important dimension for the way in which corporate governance affects firm performance.

We obtain similar patterns when we examine the impact of corporate governance on firm value and operating performance. That is, good governance is significantly positively associated with firm value and operating performance, but only among transparent firms. For opaque firms, the effect is always small and insignificant. To shed light on the channel through which good governance in transparent firms creates value, we investigate corporate investment activities. We find that firms with good governance and high transparency have less capital expenditures and engage in less acquisition activities. Considerable evidence in the literature shows a negative announcement return and a negative abnormal performance by acquiring firms. Hence these results permit the interpretation that good governance in transparent firms creates value

[5] Complementing and reinforcing our results, Mukherjee (2011) shows that shareholder rights have a positive effect on performance when shareholders possess the information needed to enforce those rights (i.e., for transparent firms).

[6] Gu (2012) includes forecast error, forecast dispersion, and revision volatility as additional predicting variables of takeover probability and finds that opacity (e.g., higher forecast dispersion) reliably reduces takeover probability.

[7] See also Hou and Robinson (2006) who find higher abnormal returns for firms in competitive industries than firms in noncompetitive industries.

by reducing agency costs.

The rest of the paper is organized as follows. The next section discusses theoretical arguments and testable hypotheses. Section 3 provides details of the data source, variable definition and summary statistics. Section 4 and section 5 examine the impact of corporate governance on stock returns. Our main results and robustness checks are, respectively, presented in these sections. Section 6 examines the relation between governance, firm value, and operating performance. Section 7 concludes.

2.2 Theoretical Arguments and Testable Hypotheses

In this section, we provide theoretical arguments for how governance and transparency can influence equity prices (or performance) and develop three competing, testable hypotheses.

To begin, the need to provide managers with incentives through governance could be smaller for more transparent firms, because outside investors more easily monitor their actions (Shleifer and Vishny (1997)). Similarly, more opaque firms lack the scrutiny of outside investors that disciplines their managers and hence they should benefit relatively more from governance. Put differently, transparency and governance are substitutes if higher transparency enhances mainly the monitoring role played by shareholders and the market, which in turn reduces managerial slack and maximizes firm performance. Since corporate governance also disciplines management, these two channels duplicate each other's positive effect on management and hence performance. Accordingly, firms with low transparency and, at the same time, fewer anti-takeover and shareholder rights provisions will, all else equal, benefit more from corporate governance than firms with high transparency.

Moreover, in less transparent environments, where managers also face riskier outcomes to their decisions and monitoring costs for outsiders are high (Demsetz and Lehn (1985)), monitoring by outsiders is relatively inefficient. Indeed, Hermalin and Weisbach (1988) find lower transparency makes the option to replace managers less valuable. If governance is a substitute for monitoring managers (i.e. reduces the likelihood of having to replace managers), then the inefficiencies due to low transparency should be lower for well-governed firms, suggesting again that governance and transparency should be substitutes in terms of their influence on firm performance. As a result, the positive effect of governance on returns should be stronger for opaque than transparent firms.

In contrast, Hermalin and Weisbach (2007) show, more disclosure need not imply higher equity value (or performance). Increased information about the firm improves the ability of outsiders to monitor their managers. However, the benefits of better monitoring do not accrue entirely to shareholders: If managers have some bargaining power, then they will capture some of these benefits via greater compensation. If better governance is a substitute for the need to provide managers with monetary incentives in case of better transparency, then their effects might offset each other. More recently, Singh and Yerramilli (2009) even establish that an increase in transparency may either increase or decrease the sensitivity of stock price to earnings, and thus, may either strengthen or weaken managerial incentives, depending on the underlying level of uncertainty. Similarly, in an extension to Paul's (1992) baseline model, he demonstrates that a higher takeover threat can actually lower real efficiency and hence lower firm value. For a detailed review of the real effects of financial markets, see Bond, Edmans, and Goldstein (2012). Overall, the impact of an increase in transparency on equity prices (or performance) is thus ambiguous. Hence one would not expect to find any reliable relation between governance, transparency, and performance.[8]

More recently, Harris and Raviv (2010) consider how allocation of control rights between shareholders and their managers interacts with a company's information environment. They invalidate claims that shareholder control reduces equity value (or performance) for opaque firms (i.e. when outsiders do not have enough information compared to insiders). Conversely, governance can improve firm performance for transparent firms.[9] Consistent with this view, we stress that more transparent firms' synergies are easier to assess by outsiders and hence these firms are more attractive takeover targets. In fact, Gu (2012) document that more transparent firms are more likely takeover targets and that this effect is statistically significant for good governance firms, but insignificant for bad governance firms. Martin and Shalev (2011) and McNichols and Stubben (2011) show that this is because acquiring firms can bid more effectively and expected synergies are larger for target firms that are more transparent. Similarly, Amel-Zadeh and Zhang (2011) and Marquardt and Zur (2011) find that low transparency (due to financial restatements or poor accrual quality) creates frictions in the market for corporate control and hence inhibits a more efficient allocation of resources via takeovers.

Consistent with this view, Duchin, Matsusaka, and Ozbas (2010) find that firm performance only increases for transparent firms when outsiders are added to the board. The interaction between governance and transparency is thus positive, because transparent firms are more likely targets, while opaque firms even might be protected from other channels that discipline their management. So, a good information environment can be crucial in facilitating takeovers and good governance only affects the performance of transparent firms. These arguments suggest that governance and transparency should be complements in terms of their influence on firm performance. Accordingly, the returns of transparent firms more reflect good governance than the ones of opaque firms.

Against the backdrop of the theoretical literature, it is an important question to evaluate which of these research directions are more in line with the data. This paper provides an empirical answer as to whether and when governance and transparency tend to influence firm performance.

2.3 Data

2.3.1 Sample Selection and Definition of Variables

Our data sources are the Investor Responsibility Research Center (IRRC), which publishes detailed governance provisions for individual firms, the Center for Research in Security Prices (CRSP), and the Institutional Brokers' Estimates System (I/B/E/S). To be included in the sample, the firm must have a match in all of these data sets. For the 1990–2006 period, this leaves us with 2,959 companies.

The IRRC tracks 24 corporate governance related provisions and the data are available for the years 1990, 1993, 1995, 1998, 2000, 2002, 2004 and 2006 during the sample period. For years when data are not

[8]Note that this subsumes non-monotonic relations, such as a u-shaped, inverted u-shape, or s-shape, which would not be detected by standard linear regression methods without appropriate conditioning (interaction) variables.

[9]See also Fishman and Hagerty (1989) for a similar argument that, by improving transparency, the firm makes it easier for outside investors to value the firm, which in turn reduces underinvestment and hence improves performance.

available, we use the observations from previous years. Based on these provisions, the Governance index ("G-index") is constructed as in GIM by adding one point to the index for the existence of each provision. The value of the G-index ranges from 0 to 24 to proxy for different degrees of corporate governance. In particular, firms with more provisions receive higher index values because they tend to have higher management power and weaker corporate governance. Firms with less provisions are assigned lower index values because they tend to have stronger shareholder rights and hence better corporate governance. Following GIM, firms with a G-index of 5 or less are referred to as democratic firms and placed into the "Democracy Portfolio." Firms with a G-index of 14 or more are referred to as dictatorship firms and placed into the "Dictatorship Portfolio." As a robustness check, we also construct the Entrenchment index ("E-index") developed in Bebchuk, Cohen, and Ferrell (2009) using the IRRC data set. The E-index is based on 6 out of 24 corporate governance provisions and the construction method is similar to that of the G-index. Firms with an E-index value of 0 fall into the "Democracy Portfolio" and firms with an E-index value of 4 or more are assigned to the "Dictatorship Portfolio."[10] As expected, the G-index and the E-index are highly correlated and the correlation between them is 0.71 over the period from 1990 to 2006.

Analysts' earning forecasts are used to gauge a firm's accounting transparency. The data on analysts' earning forecasts are obtained from I/B/E/S. Based on the analysts' earnings forecast data, we construct three transparency proxies: forecast error, forecast dispersion, and revision volatility. In particular, forecast error is defined as the absolute value of the difference between the actual annual earnings per share (EPS) and the mean of analyst forecasts. Forecast dispersion is defined as the forecast standard deviation across all analysts following the same firm in the same year. Revision volatility is computed as the standard deviation of the changes over the fiscal year in the median forecast from the preceding month. These variables are all standard in the literature and are frequently used by researchers in accounting and finance.[11] To make these measures of transparency comparable across firms, we deflate them by lagged stock price or by lagged total assets per share or by absolute value of forecast mean value. To ensure the reliability of these measures, we require that there are at least three different analysts providing forecasts for the firm during the year. All transparency proxies are constructed annually for each firm over the period from 1990 to 2006. To limit the influence of coding errors and outliers on our results, we remove observations for which forecast dispersion, forecast error, or revision volatility is larger than 10% of the share price at the beginning of the fiscal year (approximately 2% of the sample).[12] As expected, the correlation coefficients between these variables are high. Using the data for all years, the correlation between forecast error and forecast dispersion is 0.47, the correlation between forecast error and revision volatility is 0.49, and the correlation between forecast dispersion and revision volatility is 0.90.

We also use data from other sources throughout the analysis. Monthly stock return data are obtained from CRSP. Data from Compustat and CDA Spectrum are used to construct control variables for further robustness checks. Monthly observations for the three risk factors come from Kenneth French's website. The

[10]Our results are similar when we follow Bebchuk, Cohen, and Ferrell (2009), who construct Democracy portfolios using E-index scores of 5 and 6. Our cutoff is in line with other recent studies (see, e.g., Giroud and Mueller (2011)). The Dictatorship portfolio then contains sufficiently many companies relative to the Democracy portfolio.

[11]See, e.g., Givoly and Lakonishok (1979), Lang and Lundholm (1996), Thomas (2002), and Zhang (2006)

[12]See, e.g., Easterwood and Nutt (1999), Lim (2001), Teoh and Wong (2002), and Giroud and Mueller (2011).

momentum factor is constructed according to the procedure in Carhart (1997). The details will be described in later sections.

2.3.2 Empirical Relation between Transparency and Governance

Governance and transparency are potentially chosen jointly by firms and hence the evidence in the paper may not be causal. We therefore begin our analysis by examining the empirical relation between the G-index and transparency measures in a variety of ways. In essence, we find this relation is statistically insignificant, which is consistent with other studies (see, e.g., Larcker, Richardson, and Tuna (2007)). First, using all observations from 1990 to 2006, the correlation coefficients between G-index and forecast error, forecast dispersion, and revision volatility are –0.003, –0.01 and –0.01, respectively. Put differently, they are economically small. Their p–values range from 0.51 to 0.86, so none of the correlation coefficients is statistically significant.[13]

Second, we sort firms into three portfolios (i.e., lowest, medium, and highest tercile) according to their transparency measures in both Democracy and Dictatorship portfolios and find that the empirical distribution of the transparency proxies is very similar across the governance portfolios. For example, for the 1990–2006 period in Panel A of Table 2.1, firms in the lowest forecast dispersion tercile of the Democracy portfolio have the same mean and median forecast dispersion as firms in the lowest forecast dispersion tercile of the Dictatorship portfolio. Observe that the mean and median forecast dispersion is 0.001 in this tercile. Similar insights follow for the other two variables (i.e., forecast error and revision volatility) or for the sample period from 1990 to 1999 in Panel B.

Third, we investigate changes of a firm's information environment following changes in a firm's G-index and find no significant change of a firm's transparency following a change of its G-index in our sample during the 1990–1999 period. A change of the G-index for a firm in year t is computed as the difference between its G-index in year t and $t-1$. Similarly, a change of the information environment for a firm in year t is defined as the difference between the value of the transparency proxies in year t and and $t-1$. We then regress changes of a transparency proxy in year $t+1$ on changes of the G-index in year t using four samples: (1) 1,518 observations that include all firms with non-zero G-index changes; (2) 1,113 observations that only include firms with positive G-index changes; (3) 405 observations that only include firms with negative G-index changes; and (4) 215 observations that include only Democracy or Dictatorship firms with non-zero G-index changes. In these untabulated tests, we find for all transparency proxies in each of the four different samples a statistically insignificant relation between changes in governance and changes in transparency. For instance, using the fourth sample, the coefficient on the change of the G-index is –0.00092 (t-statistic = 0.54) for forecast error, it is –0.00046 (t-statistic = 0.76) for forecast dispersion, and it is –0.0015 (t-statistic = 0.76) for revision volatility.[14]

[13]Giroud and Mueller (2011) report almost no correlation between product market competition (i.e., Herfindahl index) and G-index. Yet, they find that competition and governance are substitutes in terms of firm performance. So, a substitutable or complementary effect on performance does not necessarily imply a negative or positive empirical relation between the variables of interest.

[14]The unreported results for samples (1), (2), and (3) are available from the authors upon request.

2.4 Results

2.4.1 Trading Strategies

In this section, we study the performance of trading strategies that rely on information contained in transparency proxies and in the corporate governance provisions. Recall that GIM identify a 9 percent per year disparity between the return of the Democracy portfolio and that of the Dictatorship portfolio over the period from 1990 to 1999. They employ Carhart's four-factor model to account for the style or risk differences of the two portfolios. GIM use the estimated intercept coefficient as the abnormal return to measure the effects of good governance on equity returns.

We also use Carhart's four-factor model to identify the abnormal return (i.e, the intercept α of a regression model), where the momentum factor is constructed according to the procedure in Carhart (1997). In particular, we estimate the following model:

$$R_t = \alpha + \beta_1 \times RMRF_t + \beta_2 \times SMB_t + \beta_3 \times HML_t + \beta_4 \times UMD_t + \epsilon_t \ , \tag{2.1}$$

where R_t is the excess monthly return of a common stock or a portfolio, $RMRF_t$ is the value-weighted market return minus the risk free rate in month t, SMB_t is the month t size factor, HML_t is the book-to-market factor in month t, and UMD_t is the month t Carhart momentum factor. $RMRF_t$, SMB_t, and HML_t factors are downloaded from Kenneth French's website and UMD_t factor is constructed according to Carhart (1997).

Our paper examines the joint effect of corporate governance and firm transparency on returns. Analogous to GIM and others, we construct a hedge portfolio which takes a long position in the Democracy portfolio and a short position in the Dictatorship portfolio. Similar to the procedure of Giroud and Mueller (2011) for industry competition, we split both Democracy portfolio and Dictatorship portfolio into three terciles based on the three transparency proxies.[15] This leaves us with $2 \times 3 = 6$ equal-sized portfolios for each of the three transparency proxies.[16] Therefore, we have in total $3 \times 3 = 9$ hedge portfolios with a long position in a Democracy portfolio and a short position in a Dictatorship portfolio.

To test the effect of corporate governance using the tercile portfolios based on transparency proxies, we must make sure that the distributions of the analyst variables in the same tercile are sufficiently close to each other in the Democracy and the Dictatorship portfolio. This requirement is satisfied in our sample, given the summary statistics in Table 2.1 discussed in previous subsection. In each tercile, we estimate the four-factor model and R_t is the monthly return differences between the Democracy and Dictatorship portfolios. The reported value-weighted monthly return is calculated by weighting the return of each individual stock in the portfolio by its market capitalization at the end of previous month and equal-weighted monthly return is the average of the return of each individual stock in the portfolio.

The IRRC updates corporate provisions in August 1990, June 1993, June 1995, January 1998. We assign new values of G-index to firms one month after the IRRC updates. So the Democracy and Dictatorship

[15] In unreported results, we have experimented with the sample median of the transparency proxies to create two (instead of three) groups and quartile portfolios to create four (instead of three) groups. Both methods lead to similar findings.

[16] We have verified that our results are invariant to the order in which we construct portfolios. Recall that the correlation between the transparency proxies and the G-index is essentially zero and insignificant.

portfolios are reset in September 1990, July 1993, July 1995 and February 1998. The transparency measures are calculated annually using yearly I/B/E/S data. To avoid the look-ahead bias, we rebalance the hedge portfolio each July using previous year's value of transparency proxies. Our main analysis focuses on the period from September 1990 to December 1999, which is the sample period used in GIM and others. In robustness checks we also extend the sample period to 2011.

2.4.2 Baseline Results

In Table 2.2, we report GIM's original results and our replication of their results. The first row in Panel A is the result in GIM's (2003) paper. It shows that the value-weighted Democracy-Dictatorship hedge portfolio earns a monthly abnormal return of 0.71 percent, which is statistically significant at the 1% level. Row (2) reports our replication of their result. Observe that our estimate of α equals 0.67 percent, with significance at the 1% level. The alpha and factor loadings are very similar, but not identical.[17] In Row (3), we perform the same estimation using our restricted sample and obtain a reliably positive α of 0.68% (t-statistic = 2.57) as a base case. The coefficients for the risk factors are similar to those reported in GIM, who only document results for value-weighted portfolios. However, for comparison to our results based on equal-weighted portfolios, we report in the second row of Panel B a replication of GIM's (2001) equal-weighted results, which are presented in the first row. As revealed by the table, the alpha and factor loadings are again very similar in terms of economic magnitude and statistical significance. Finally, the third row tabulates the baseline results for our restricted sample.

Table 2.3 presents our main results. We divide both Democracy and Dictatorship portfolios into 3 equal-sized (tercile) portfolios based on the three transparency proxies: forecast dispersion, forecast error, and revision volatility. This leaves us with three Democracy-Dictatorship hedge portfolios. We obtain the abnormal return α by running a time series regression of the monthly excess returns of each hedge portfolio on the market factor ($RMRF$), the size factor (SMB), the book-to-market factor (HML), and the momentum factor (UMD). Panel A reports the abnormal returns of the trading strategy when transparency proxies are scaled by lagged share price. We provide results for both equal-weighted and value-weighted portfolios. In the first row of Panel A, forecast dispersion is used to construct the terciles. The estimated abnormal return is reliably positive for the hedge portfolio in the lowest tercile, while the alphas in the medium and highest tercile are much smaller and insignificant or less significant. Notably, this monotonically declining pattern for alpha prevails for both equal-weighted and value-weighted portfolios. More specifically, for the equal-weighted portfolios, the alpha is 0.73% and significant at the 5% level (t-statistic = 2.41) in the lowest tercile, it is 0.40% and not significant (t-statistic = 1.31) in the medium tercile, and it is 0.05% without significance (t-statistic = 0.18) in the highest tercile. The difference between the alphas in the lowest and highest terciles equals 0.68%, which is significant at the 10% level. For the value-weighted portfolios, the alpha is 0.99% with a significance at 1% level in the lowest tercile (t-statistic = 3.00), while it is economically smaller and statistically insignificant in the medium and the highest terciles. The difference between the alphas in the lowest and highest terciles equals 1.20%, which is significant at the 5% level. Finally, notice that the economic

[17] For example, Core, Guay, and Rusticus (2006) and Giroud and Mueller (2011) report slightly different estimates for their replications, which are more in line with our replication.

magnitude of the abnormal return in the lowest tercile exceeds the one in Table 2.2 for the full sample.

In the second and third row of Panel A, we report the abnormal returns of the trading strategy when analyst forecast error and revision volatility are employed, respectively. Interestingly, similar abnormal return patterns obtain when replacing forecast dispersion by these variables. That is, we always estimate a positive and significant abnormal return in the lowest tercile and we get small and insignificant abnormal returns for most of the estimations in other terciles. For example, for the equal-weighted portfolios formed based on forecast error and corporate governance, the abnormal return equals 0.80% with 1% significance (t-statistic = 2.88) in the lowest tercile and 0.01% with no significance (t-statistic = 0.01) in the highest tercile. Using revision volatility, the alpha is 0.52% and significant (t-statistic = 2.02) in the lowest tercile of the equal-weighted hedge portfolio, while it is 0.13% and insignificant (t-statistic = 0.41) in the highest tercile.

The results show that the Democracy-Dictatorship trading strategy is more effective when it is restricted to firms for which analysts have less difficulties in making forecasts. Since firms' transparency facilitate actual takeovers, the positive effect of lack of anti-takeover and shareholder rights provisions (good governance) is strengthened. GIM's sample draws on firms with various levels of transparency. Clearly, their estimates of abnormal returns are an average effect that describes the relation between governance and returns for both opaque and transparent firms. Recall that, for example, the equal-weighted abnormal return is 0.45% in GIM's (2001) study. However, the abnormal return in our lowest terciles is much larger, ranging from 0.52% to 0.80%, because these lowest terciles restrict the trading strategy to include only the more transparent firms from the full sample.

Panel B of Table 2.3 contains the estimation results for the same tests as in Panel A except that the transparency proxies are scaled by lagged total assets per share instead of lagged stock price. Notably, the transparency-related patterns of abnormal returns are qualitatively and quantitatively very similar to the ones in the previous table. Observe, for example, that the value-weighted alpha of the trading strategy based on forecast dispersion is 0.79% and significant (t-statistic = 2.32) in the lowest tercile, it is 0.36% and insignificant (t-statistic = 0.81) in the medium tercile, and it is 0.29% and insignificant (t-statistic = 0.40) in the highest tercile. Thus, irrespective of the deflator for transparency proxies, we find a monotonically declining pattern in abnormal returns when forecast error, forecast dispersion, or revision volatility rises (i.e., transparency deteriorates).

Panel C of Table 2.3 reports results when transparency measures are scaled by absolute value of forecast mean value instead of lagged stock price or lagged assets per share.[18] Again we obtain similar patterns for abnormal return alpha. For instance, the value-weighted alpha of the trading strategy based on forecast error is 0.83% and significant (t-statistic = 2.55) in the lowest tercile, it is 0.67% and less significant (t-statistic = 1.93) in the medium tercile, and it is –0.04% and insignificant (t-statistic = 0.08) in the highest tercile.

To see which side of the hedge portfolio contributes to its abnormal returns in the lowest tercile, we perform return decomposition analysis. The last two columns of Table 2.3 report the value-weighted four-factor α of the Democracy and Dictatorship portfolios in the high transparency groups (i.e., lowest terciles). Observe that the long leg of the hedge portfolio is associated with significantly positive abnormal returns,

[18] We have also verified in unreported estimations that deleting all observations with negative forecast mean does not affect our results.

while the α of the short leg is negative, economically small, and not statistically significant. For example, in Panel A, when the transparency proxy is forecast dispersion, the long-side and short-side αs are, respectively, 0.81% (t-statistic = 3.92) and –0.18% (t-statistic = 0.79), resulting in a 0.99% abnormal return to the hedge protfolio. Furthermore, this pattern holds in all cases across all three panels.[19] Overall, the result of this analysis indicates that the abnormal returns to the hedge portfolio largely comes from the long side and good governance firms are associated with positive and significant abnormal returns among transparent firms.

2.4.3 Time-Series Average of Transparency Proxies

We also perform our baseline tests with time-series averages of forecast dispersion, forecast error and revision volatility, which are largely outside of managers' discretion and hence more *permanent* firm characteristics. These time-series averages are time-invariant and hence more exogenous. Hence the hedge portfolios are not affected by potentially strategic information disclosure or time-varying information arrival. Table 2.4 collects the results when we use *time-invariant* sample averages of transparency proxies for three different deflators.[20] We report abnormal returns for both equal-weighted and value-weighted hedge portfolios using the Carhart four-factor model. The prevailing pattern in the table reinforces our earlier findings. In all tests, we find large and significant abnormal returns in the lowest tercile, which contains transparent firms, and small and insignificant abnormal returns for the highest tercile, which contains opaque firms. Perhaps surprisingly, when we remove the time-series variation of the transparency measures, the economic and the statistical significance of the abnormal returns in the high transparency terciles remain largely unchanged.

2.4.4 Principal Component of Transparency Proxies

So far we use the three transparency proxies individually to proxy for firm transparency level. Although the three variables are correlated with each other, they still convey different characteristics of analyst earnings forecasts and thus contain different information about the firm's information environment. In this section, we combine all the information contained in forecast error, forecast dispersion, and revision volatility by constructing their first principal component. Principal component analysis is appropriate to reduce several observed variables into a smaller number of principal components that account for most of the variation in the observed variables. That is, we replace the transparency measures with their first principal component and re-estimate the four-factor model in Equation (1). The estimation results are presented in Table 2.5 for equal-weighted and value-weighted Democracy-Dictatorship hedge portfolios.

Panel A of Table 2.5 reports the alphas for the trading strategy when transparency proxies are scaled by lagged share price. For the equal-weighted hedge portfolios, the alpha is 1.28% and significant (t-statistic = 3.50) in the lowest tercile, it is –0.22% and insignificant (t-statistic = 0.69) in the medium tercile, and it is 0.04% and insignificant (t-statistic = 0.11) in the highest tercile. For the value-weighted hedge portfolios,

[19] We also perform the return decomposition analysis to the equal-weighted portfolios and obtain similar results. For instance, in the case of forecast dispersion scaled by lagged price, the long-side and short-side α is 0.70% (t-statistic = 3.71) and –0.04% (t-statistic = 0.15) respectively. We do not tabulate the equal-weighted return decomposition results in Table 2.3 (to save space), but they are available from the authors upon request.

[20] In untabulated results, we also use the sample averages of transparency proxies over the period of 1982 to 1990 as a time-invariant measure of a firm's transparency level and obtain similar abnormal return patterns.

the alpha is 1.37% and significant (t-statistic = 3.52) in the lowest tercile, it is 0.15% and insignificant (t-statistic = 0.31) in the medium tercile, and it is 0.04% and insignificant (t-statistic = 0.09) in the highest tercile. The relation between abnormal returns and firm transparency we obtain in this table again strongly supports our previous findings. In addition, notice that combining the individual information from the three variables yields economically larger estimates for alpha. More specifically, Panel A shows that the value-weighted hedge portfolio — provided it only includes more transparent firms as defined by the first principal component of forecast error, forecast dispersion, and revision volatility — earns a monthly alpha of 1.37% (t-statistic = 3.52), which is almost twice as large as the alpha estimate of 0.71% reported in GIM.

In Panels B and C, we employ the transparency proxies' first principal component scaled by lagged total assets per share and absolute value of forecast mean, respectively, and re-estimate the four-factor model in Equation (1). The test results are similar to the ones reported in Panel A.

Overall, Tables 2.3 to 2.5 provide novel and strong support for the hypothesis that more transparent firms benefit relatively more from good corporate governance, while more opaque firms benefit relatively less from good governance. Because for transparent firms, there is more precise information available to investors, it is less difficult for outsiders to assess synergies, which facilitates actual takeovers (see Amel-Zadeh and Zhang (2011), Marquardt and Zur (2011), Martin and Shalev (2011), and McNichols and Stubben (2011)). This strengthens the effect of good governance. Thus, the abnormal returns of the trading strategy focusing on transparent firms are positive and significant. On the other hand, there is less precise information available for opaque firms, which we proxy by the noise contained in analyst earnings forecasts. As a result, it is more difficult, if not impossible, for outsiders to assess synergies. Thus opaque firms might not be taken over even if they lack anti-takeover and shareholder rights provisions. So the abnormal returns of the trading strategy focusing on opaque firms are small and insignificant.

Consistent with our main result, we find that target firms are, on average, relatively more transparent. We compare the average forecast dispersion, forecast error, or revision volatility of target firms to all firms in the sample over the 1990–2006 period. For example, the average forecast dispersion is 0.043 for the subsample of target firms, while it is 0.197 for the sample of all firms, and the difference between the two is statistically significant at the 1% level. Moreover, Gu (2012) employs an empirical logit model including transparency as additional predicting variable. She documents that firm transparency has a significantly positive association with takeover likelihood. Using this specification, we find in untabulated results that only good governance firms, which are transparent, are more likely takeover targets. Taken together, these results support the view that transparency facilitates takeovers and hence increases takeover vulnerability.

2.5 Robustness

2.5.1 Variants of the Trading Strategy

Table 2.6 presents the results for the first set of robustness checks. In Panel A, accruals quality is used to measure firm transparency instead of analyst variables. Following McNichols (2002),[21] accruals quality is

[21] See also Dechow and Dichev (2002), Francis, LaFond, Olsson, and Schipper (2005), McNichols and Stubben (2011).

constructed as the standard deviation of the residuals of the following estimation model:

$$\Delta WC_t = b_0 + b_1 CFO_{t-1} + b_2 CFO_t + b_3 CFO_{t+1} + b_4 \Delta Sales_t + b_5 PPE_t + \epsilon_t \ , \tag{2.2}$$

where ΔWC_t is the change in working capital from year $t-1$ to year t. Specifically, it is computed as the increase in accounts receivable (Compustat item #302) plus the increase in inventory (item #303) minus the increase in accounts payable and accrued liabilities (item #304) minus the increase in taxes accrued (item #305) plus the increase (decrease) in other assets or liabilities (item #307). CFO is operating cash flow (item #308), $\Delta Sales_t$ is change in sales (item #12) from year $t-1$ to year t, and PPE is property, plant, and equipment (item #8). All variables are scaled by lagged total assets. Each year this model is estimated for every firm using prior eight years and the standard deviation of the residual is defined as the accruals quality.[22] A larger standard deviation means lower accruals quality and lower accounting transparency. As is shown in Panel A, this alternative transparency measure not only confirms our main findings but also produces stronger results. For example, the value-weighted alpha in the lowest tercile is 1.15% (t-statistic = 2.97), which is larger than the corresponding value-weighted alpha in Table 2.3.

In Panel B, we use E-index instead of G-index as an alternative measure of corporate governance. Bebchuk, Cohen and Ferrell (2009) propose the E-index, which they base on the six most important provisions among the 24 of the G-index. They suggest that these six provisions are the key determinants of GIM's findings and hence a less noisy measure of corporate governance. Following Giroud and Mueller (2011), firms with E-index value of 0 are assigned to "Democracy Portfolio" and firms with E-index value of 4 or more are assigned to "Dictatorship Portfolio". As is shown in the panel, the pattern of the alpha is very similar to what we obtain when we use the G-index. For instance, in the case of forecast error, the alpha estimated using equal-weighted portfolios is 0.83% and highly significant (t-statistic = 4.79) in the lowest tercile, alpha is 0.34% and less significant (t-statistic = 1.75) in the medium tercile, and it is 0.13% with no significance (t-statistic = 0.61) in the highest tercile. Observe that the pattern is not only similar for value-weighted portfolios but also for the estimation results involving the other two transparency proxies.[23]

Panel C and Panel D of Table 2.6 report abnormal returns for the other two robustness checks that rely again on the G-index. In Panel C, we exclude firms that belong to the "new economy" group defined in Hand (2003). This group contains 274 Internet stocks. Core, Guay, and Rusticus (2006) show that the value-weighted abnormal return of the Democracy-Dictatorship hedge portfolio drops from 0.71% (t-statistic = 2.73) to 0.46% with decreased significance (t-statistic = 1.83). So the large abnormal returns might be driven by those new economy internet firms. We check if similar patterns prevail in the sample without those firms. As is shown in Panel B, most of the alphas are large and highly significant in the lowest tercile and all alphas are small and insignificant or marginally significant in the medium and highest terciles. Take forecast error as an example, the equal-weighted alpha is 0.72% and significant (t-statistic = 2.65) in the lowest tercile, it is 0.21% and insignificant (t-statistic = 0.75) in the medium tercile, and it is –0.04% and insignificant (t-statistic = 0.12) in the highest tercile. Similar findings obtain for value-weighted hedge

[22]We also estimate the model using prior ten years or twelve years and the results do not change qualitatively.

[23]In unreported tests, we find qualitatively identical results using only a (0, 1, 2) count variable for poison pills and classified boards instead of the G-index. These provisions are commonly perceived as the two most effective takeover defenses and hence reinforce the takeover story. We thank the anonymous referee for suggesting this line of analysis.

portfolios. However, comparing to our main results using the full sample, the magnitude and significance of alpha drops especially in case of value-weighted portfolios, which is consistent with the results reported by Core, Guay, and Rusticus (2006).

In Panel D of Table 2.6, we extend the sample period to 2011. Prior studies find that the alpha drops when the sample period is extended. For example, Core, Guay, and Rusticus (2006) estimate a decrease to 0.40% in alpha (t-statistic = 1.68) when the sample period is extended to December 2003. Giroud and Mueller (2011) document that the value-weighted alpha drops to 0.24% and the equal-weighted alpha drops to 0.29% when their sample is extended to December 2006. The alpha of our extended sample drops to 0.23% (t-statistic = 1.38) for the value-weighted hedge portfolio and to 0.32% (t-statistic = 2.01) for the equal-weighted hedge portfolio (not reported). Even though the results for the extended sample are weaker, we still find similar abnormal return patterns when we divide the Democracy and Dictatorship portfolios into terciles based on the transparency proxies in some cases. Finally, Panel E of the table shows results for the 2000–2011 period. Consistent with the recent literature (see, e.g., Bebchuk, Cohen, and Wang (2011)), the results disappear in the 2000s.

Table 2.7 continues our robustness checks by investigating whether transparency is just correlated with other factors (e.g., institutional ownership or industry concentration) that are already known to be important. that is, we gather estimation results for these subsample tests, which help rule out such an omitted variable story by showing that our result holds within these different subsamples. Put differently, this would suggest that transparency is independently important.

In Panels A and B of Table 2.7, we split the sample into firms with high and low institutional ownership, respectively. Institutional ownership is defined as the percentage of shares held by the 18 largest public pension funds listed in the Appendix of Cremers and Nair (2005).[24] We first divide Democracy and Dictatorship portfolios into equal-sized subgroups with high and low institutional ownership based on the level of firm's 13F holdings data. The cutoff is the median of the distribution. This leaves us with 4 portfolios. Then we split each of these 4 portfolios into 3 equal-sized subgroups based on each of the three transparency proxies. For each transparency measure, we end up with 12 portfolios: 6 with high institutional ownership and 6 with low institutional ownership. We again form hedge portfolios and estimate Cahart's four-factor model within these subgroups.

Cremers and Nair (2005) find that the Democracy-Dictatorship portfolio earns significant abnormal return only for the group including firms with high institutional ownership. Our results are consistent with their findings in both cases of value-weighted portfolios and equal-weighted portfolios. In the case of forecast error, for example, the value-weighted alpha is 1.41% with a t-statistic of 3.67 in the lowest tercile of the group with high institutional ownership, while it is 0.56% with a t-statistic of 1.14 in the lowest tercile of the group with low institutional ownership. In the case of forecast dispersion, the value-weighted alpha is 1.54% with a t-statistic of 3.53 in the lowest tercile of the group with high institutional ownership, while it is 0.27% with a t-statistic of 0.61 in the lowest tercile of the group with low institutional ownership. Similar insights arise in the case of revision volatility. More importantly, these results are confirming our main findings. For the high-institutional-ownership group, we obtain large and significant estimates for alpha in the lowest

[24]For these robustness tests, we retrieve and use the CDA Spectrum Institutional (13F) Holdings data.

tercile and smaller and insignificant or marginally significant alphas in the medium and highest terciles. This is true for all three transparency proxies. For example, in the case of forecast error, the trading strategy in the lowest tercile earns an equal-weighted abnormal return of 1.08% at the 1% significance level (t-statistic = 3.11), while the alpha in the medium tercile is 0.34% with no significance (t–statistic = 0.95) and the estimated abnormal return in the highest tercile is 0.24% with no significance (t–statistic = 0.59).

In Panels C and D of Table 2.7, we split the sample into firms in industries with low and high product market competition to verify that our findings are not driven by product market competition. Giroud and Mueller (2011) establish significant abnormal returns of the Democracy-Dictatorship portfolio in the subgroup that includes firms in less competitive industries and conclude that firms in concentrated industries benefit more from good corporate governance. To examine whether our results are largely due to industry competition, we employ a sales-based Herfindahl-Hirschman index ("HHI") to proxy for industry competition. Following Giroud and Mueller (2011), we compute HHI as the sum of squared market shares for firms in each of the 48 Fama and French (1997, FF) industries. Market shares are computed using firms' sales data from Compustat. We first divide Democracy and Dictatorship portfolios into two equal-sized subgroups with high and low industry competition based on the level of the HHI. This leaves us with 4 portfolios. Then we divide each of these 4 portfolios into 3 equal-sized subgroups based on the three transparency proxies. Thus, we end up with 12 portfolios: 6 with low industry competition and 6 with high industry competition. We again form the Democracy-Dictatorship hedge portfolios and estimate the four-factor model within the low-industry-competition and the high-industry-competition subgroups.

Panel C of Table 2.7 presents the estimated alphas for the portfolios formed in concentrated industries (i.e., industries with lower competition or higher values of HHI). Again, our estimates of abnormal returns show similar patterns as in our main results. Alpha is always large and significant in the lowest tercile and smaller and insignificant or marginally significant in the medium and highest tercile. For instance, in the case of forecast dispersion, the equal-weighted alpha is 0.95% and significant (t-statistic = 2.53) in the lowest tercile, and it is 0.39% and insignificant (t-statistic = 0.05) in the highest tercile. Perhaps surprisingly, the value-weighted alpha is 1.97% and significant (t-statistic = 3.68) in the lowest tercile and it is 0.46% and insignificant (t-statistic = 0.56) in the highest tercile. Panel D presents the estimated alphas for the portfolios formed in competitive industries (i.e., industries with higher competition or lower values of HHI). In this subsample, our baseline findings are still valid, but, as expected, economic magnitudes and statistical significance levels are smaller than in concentrated industries. Consistent with Giroud and Mueller (2011), the transparency-related patterns of abnormal returns are diminished in Panel D relative to Panel C. Intuitively, a higher level of product market competition is a substitute for governance and hence weakens the effect of corporate governance on equity prices. In sum, this set of robustness tests also supports our main findings.

Table 2.8 presents the results for the final set of robustness checks. In this table, we study how our findings are affected by firm characteristics, such as financial leverage and firm size, which could be correlated with transparency and hence explain our result. In Panels A and B, we split the sample into firms with a low (below-median) leverage ratio and a high (above-median) leverage ratio, respectively. It has been argued that high leverage reduces the probability of a takeover.[25] Thus, firms with high levels of debt are more

[25]See, e.g., Stulz (1988), Novaes and Zingales (1995), Zwiebel (1996), Harris and Raviv (1998).

difficult to be taken over, even though they may have fewer anti-takeover and shareholder rights provisions. This suggests that corporate governance should be more effective among low leverage firms. Comparison of the results in Panel A and B provides support for this view. For example, in case of forecast error, the value-weighted alpha is 1.67% with a t-statistic of 4.51 in the lowest tercile of the low-leverage group, while it is 1.05% with a t-statistic of 2.41 in the lowest tercile of the high-leverage group. In case of forecast dispersion, the value-weighted alpha is 1.58% with a t-statistic of 3.53 in the lowest tercile of the group with low leverage, while it is 1.13% with a t-statistic of 2.84 in the lowest tercile of the group with high leverage. Similar results arise for revision volatility. More importantly, these results from both panels confirm our main findings. In all cases and for *both* groups, we obtain large and significant alphas in the lowest tercile and smaller and insignificant alphas in the highest tercile.

In Panels C and D of Table 2.8, we split the sample into firms with low (below-median) asset value and to firms with high (above-median) asset value, respectively. Motivated by the view that large (target) firms require bidders to spend more effort and resources and thus firm size has a deterrent role for takeovers, governance mechanisms should be more effective for smaller firms. The results in Panels C and D confirm this view. The alphas in the lowest tercile in all cases in Panel C are economically larger than the corresponding alphas in the lowest tercile in Panel D. For instance, in the case of forecast error, the value-weighted alpha is 1.59% with a t-statistic of 2.60 in the lowest tercile of the small-firm group, while it is 0.92% with a t-statistic of 2.58 in the lowest tercile of the large-firm group. In sum, this set of subsample studies not only confirms our main findings, but also provides support for the idea that the effectiveness of the governance mechanisms can be affected by factors that have an impact on the takeover probability. That is, transparency matters more for smaller firms, but again these tests suggest that transparency is independently important, because it also matters for larger firms.

2.5.2 Industry Effects

In this section, we adjust for industry effects. It is very important because some industries tend to have high abnormal returns and some tend to have low abnormal returns (see, e.g., Hou and Robinson (2006)). If a large part of the Democracy portfolio is formed by firms from high abnormal return industries and a large part of the Dictatorship portfolio is formed by firms from low abnormal return industries, this could be one possible source of the large and significant abnormal returns in our findings. Thus, we re-estimate the regression using industry-adjusted returns, which are obtained by subtracting the median monthly industry return from the individual firm's monthly return. The median industry return is calculated in each of the 48 FF industries using all firms in the CRSP universe. These estimation results are presented in Table 2.9.

As is shown by the table, adjusting for industry returns does not weaken our previous findings. That is, similar patterns appear again for all of the three modified regression models. In Panel B, for example, we estimate the Carhart 4-factor model with industry-adjusted returns. We again find large and significant alphas in the lowest tercile, smaller and insignificant or marginally significant returns in the medium and highest terciles for both equal-weighted and value-weighted portfolios. Notice that, in the case of forecast error and equal-weighted returns, the alpha is 0.81% and significant (t-statistic = 3.35) in the lowest tercile, it is 0.19% and insignificant (t-statistic = 0.84) in the medium tercile, and it is –0.05% and insignificant

(t-statistic = 0.21) in the highest tercile.

2.5.3 Fama-MacBeth Return Regressions

In this section, we follow the literature by testing our hypothesis using standard Fama-MacBeth return regressions with a set of control variables and find supportive evidence for our main result. In order to identify the effect of corporate governance on stock returns in different transparency groups, we create a dummy variable for the tercile of firm's transparency and interact it with the Democracy dummy. So, we estimate the following model:

$$R_{it} = \alpha_t + \beta_{1t}(D_{it} \times A1_{it}) + \beta_{2t}(D_{it} \times A2_{it}) + \beta_{3t}(D_{it} \times A3_{it}) + \gamma_t \mathbf{Z}_{it} + \epsilon_{it} \ , \tag{2.3}$$

where R_{it} is the month t stock return of firm i, D_{it} is the Democracy dummy (which equals to one if the firm is in the Democracy portfolio and equals to zero if the firm is in the Dictatorship portfolio), $A1_{it}$ is a dummy variable for the lowest tercile of the analyst variable, $A2_{it}$ is a dummy variable for the medium tercile, $A3_{it}$ is a dummy variable for the highest tercile, and $\mathbf{Z}_{it}$ is a vector of control variables. Following GIM, the elements of $\mathbf{Z}_{it}$ include book-to-market ratio, gross return from month $t-3$ to month $t-2$, gross return from month $t-6$ to month $t-4$, gross return from month $t-12$ to month $t-7$, firm size, leverage, stock price, sales growth over previous five years, trading volume of NYSE or Amex stocks, trading volume of NASDAQ stocks, a NASDAQ dummy, an S&P 500 dummy, dividend yield, institutional ownership, product market competition (measured by a sales-based Herfindahl-Hirschman index), and firm idiosyncratic volatility. All explanatory variables are lagged. The control variables also include dummies of the analyst variables to account for the direct effect of the transparency terciles. We control for firm idiosyncratic volatility is because Ferreira and Laux (2007) argue that firms with lower level of G-index show higher idiosyncratic volatility. We estimate Equation (2.3) every month over the period from September 1990 to December 1999 and report the mean of the monthly estimate of the coefficients for relevant variables.[26]

Table 2.10 summarizes the results. In column [1], G-index is used as the measure of corporate governance and a sample with all firms is used in the monthly cross-sectional regression. As is shown, the coefficient on G-index is –0.03. Although the sign is what we expect to show the positive impact of good governance on stock returns, the magnitude is small and the statistical significance is low (t-statistic = 1.45). This is consistent with the results in GIM. In column [2], we restrict the sample to firms in the Democracy and Dictatorship portfolios and use a Democracy dummy to proxy for corporate governance. This is consistent with the way the Democracy-Dictatorship portfolio is constructed in the prior analysis. As displayed in the table, the coefficient estimate on the Democracy dummy is 0.34 but insignificant (t-statistic = 1.51). It means that, on average, a Democracy firm can earn a 0.34% higher monthly return than a Dictatorship firm, if all other firm characteristics are the same. In column [3], the Democracy dummy is interacted with three transparency dummies (using forecast dispersion) to identify the differential effect of corporate governance

[26]The idea of the regression specification here is to estimate the impact of corporate governance on stock returns across different transparency groups. Thus it is not necessary to include an extra term for the G-index dummy to identify the governance effect again. In order to control for any direct effect from transparency on stock returns, we also include dummy variables for high and medium transparency terciles. The low transparency tercile's dummy variable is not included because of redundancy—a similar specification is employed in Giroud and Mueller (2011).

on stock returns for different transparency groups. The results also support our main finding. That is, the coefficient on the interaction term for the lowest tercile is 0.64 and significant (t-statistic = 2.43), the coefficient on the interaction term for the medium tercile is 0.27 and insignificant (t-statistic = 1.10), and the coefficient on the interaction term for the highest tercile is –0.43 and insignificant (t-statistic = 1.25). Similar patterns prevail in columns [4] and [5], where we replace forecast dispersion by forecast error and revision volatility, respectively. Taken together, these Fama-MacBeth regression results reinforce our baseline results.

2.5.4 Alternative Asset Pricing Models

In this section, we further estimate the abnormal returns of the trading strategy using alternative asset pricing models. The results are reported in Table 2.11. In Panel A, we replace the Carhart four-factor model with the market (or capital asset pricing model) model. As is shown in the table, the alphas for the transparent group are positive and significant for both equal-weighted and value-weighted portfolios, while the alphas for the opaque group are small and insignificant. This is consistent with our main hypothesis. However, the results are weaker in terms of both magnitude and statistical significance. This is consistent with Giroud and Mueller (2011), who also find weaker results when using market model, perhaps because the Democracy-Dictatorship hedge portfolios have significant and negative loadings on both size and value factor. As seen in Panel A of Table 2.2, both GIM and our replication results reveal that the hedge portfolios load negatively on size and value factor, so excluding these factors in the regression model will result in a less positive intercept.

In Panel B, we use the Fama-French four-factor model to replace the Carhart four-factor model. The two models are almost the same. They share the same market factor, size factor, and book-to-market factor, but have different momentum factors. Fama-French momentum factor is constructed by using double sorting on firm size and stock momentum, while Carhart momentum factor is constructed by sorting stocks on momentum only.[27] According to Panel B, this factor construction difference does not affect our results that much. We still obtain similar abnormal return patterns across the three transparency groups, although the magnitude of the alphas is smaller.

In Panel C, we augment the Carhart four-factor model with the liquidity factor of Pastor and Stambaugh (2003), who find that market-wide liquidity is an important asset pricing variable. As is shown in the table, adding this additional risk factor results in economically and statistically similar abnormal return patterns.

In Panel D, we extend the Carhart four-factor model by the takeover factor of Cremers, Nair, and John (2009), who construct this factor based on a target firm's takeover probability.[28] Cremers, Nair, and John (2009) find that the abnormal returns to the Democracy-Dictatorship portfolio in GIM become insignificant when they include the takeover factor in the four-factor model. However, adding this additional factor does not fully explain our abnormal return patterns across transparency terciles. For example, the value-weighted alpha decreases from 1.10% (t-statistic = 2.45) using Carhart four-factor model (see Panel B of Table 2.5 for this estimation result) to 0.92% (t-statistic = 2.16) using the extended five-factor model, suggesting that the takeover factor of Cremers, Nair, and John (2009) does explain part of the abnormal returns to the Democracy-Dictatorship portfolios. Since the alpha for the high transparency tercile remains positive

[27] The data for their momentum factor are obtained from Kenneth French's website.

[28] We are grateful to Martijn Cremers for providing us with the takeover factor data.

and significant for both equal-weighted and value-weighted portfolios, we conclude that a firm's information environment creates another important dimension that matters for the way in which a firm's governance mechanism affects firm performance. Put differently, it is likely that a firm's information environment has an incremental impact on its takeover likelihood (see, e.g., Amel-Zadeh and Zhang (2011), Marquardt and Zur (2011), Martin and Shalev (2011), and McNichols and Stubben (2011)). Therefore, transparency might be an additional variable for gauging a firm's takeover probability, which might significantly enhance the takeover factor of Cremers, Nair, and John (2009).

2.6 Governance, Firm Value, and Operating Performance

To provide further evidence, we follow GIM and study the effect of corporate governance on firm value (i.e., Tobin's Q) and on measures of operating performance. While GIM examine these relations for all firms, Giroud and Mueller (2011) find that they are more pronounced for firms in noncompetitive industries. Here we investigate whether the positive role of governance varies across firms with different transparency levels.

2.6.1 Governance, Transparency, and Firm Value

To study the relation between governance and firm value, we estimate the following model:

$$Q_{it} = \alpha_t + \alpha_j + \beta_{1t}(G_{it} \times A1_{it}) + \beta_{2t}(G_{it} \times A2_{it}) + \beta_{3t}(G_{it} \times A3_{it}) + \gamma_t \mathbf{Z}_{it} + \epsilon_{it}, \quad (2.4)$$

where Q_{it} is the industry-adjusted Tobin's Q in year t for firm i, α_j and α_t are industry and year fixed effects, G_{it} is the G-index, $A1_{it}$ is the dummy variable for the lowest tercile of analyst variable, $A2_{it}$ is the dummy variable for the medium tercile of analyst variable, $A3_{it}$ is the dummy variable for the highest tercile of analyst variable, Z_{it} is a set of control variables including firm age (in logs), firm size (book value of assets, Compustat item #6, in logs), S&P 500 dummy and a Delaware dummy. These are also the control variables used in GIM (2003). The control variables also include dummies of the analyst variables to account for the direct effect of the transparency terciles. Tobin's Q is defined as the market value of assets divided by the book value of assets (item #6). Market value of assets is calculated by using the sum of book value of assets (item #6) and market value of common equity from Compustat (item #24 × item #25) minus the sum of book value of common equity (item #60) and balance sheet deferred taxes (item #74). Industry-adjusted Tobin's Q is calculated by adjusting firm's Q by the industry median. Industry median is computed every year for each of the 48 FF industries. The sample period is from 1990 to 2006 and the standard errors are clustered at the industry level.

Table 2.12 contains our results for the panel regression. Column [1] reports the coefficient estimate from the regression of Tobin's Q on G-index. The coefficient is –0.035 and highly significant (t-statistic = 2.75), confirming the positive impact of corporate governance on firm value identified in the literature. In column [2], we interact the G-index with forecast error dummies and report the coefficients for the interaction terms. Similar patterns as in our earlier tables appear: the absolute value of the coefficient is largest (–0.095) and most significant (t-statistic = 3.35) in the transparent group, is smaller (–0.023) and less significant (t-statistic = 2.60) in the medium group, and is smallest (0.008) and insignificant (t-statistic = 1.06) in the opaque group. In columns [3] and [4], we interact the G-index with forecast dispersion and revision volatility dummies, respectively. We also obtain monotonic patterns for these two transparency measures.

2.6.2 Governance, Transparency, and Operating Performance

To study the relation between governance and operating performance, we use the same regression specification as in Equation (4) but replace industry-adjusted Tobin's Q by measures of firm performance: return on assets (ROA), return on equity (ROE), and net profit margin (NPM). ROA is defined as net income (item #172) divided by the book value of assets (item #6), ROE is defined as net income divided by the book value of common equity (item #60), and NPM is defined as net income divided by sales (item #12). All performance variables are industry-adjusted by subtracting the industry median from the performance variable. Industry median is computed every year for each of the 48 FF industries. All dependent variables are trimmed at the 5th and 95th percentiles of their empirical distribution. All explanatory variables are lagged. Following GIM, we include the logarithm of the book-to-market ratio in the previous year as an additional control variable. The sample period is from 1990 to 2006 and the standard errors are clustered at the industry level.

Table 2.13 contains our results for the panel regression. Columns [1] to [3] report the coefficients on the interaction terms when the dependent variable is return on assets (ROA) and transparency is measured by forecast error, forecast dispersion, and revision volatility, respectively. The coefficient on the interaction term for the dummy variable of the lowest tercile is always negative and statistically significant, while the coefficient on the interaction term for the dummy variable of the highest tercile is small and insignificant. For instance, in column [1], the coefficient estimate for the G-index is large (–0.182) and significant (t-statistic = 2.28) for the transparent group, and it is small (0.007) and insignificant (t-statistic = 0.13) for the opaque group. In columns [4] to [6], the dependent variable is net profit margin (NPM), and in columns [7] to [9], the dependent variable is return on equity (ROE). In both cases, we find similar patterns for the coefficient estimates of all three transparency measures. That is, the coefficient of interest is negative and significant for transparent firms, and it is small and insignificant for opaque firms. This confirms the earlier finding that firms with better governance have better operating performance, and also reveals that the positive effect of governance is largely coming from transparent firms.

2.6.3 Capital Expenditure and Acquisition Activity

The results from our firm value and operating performance tests indicate that good governance firms are associated with higher firm value and higher operating performance. To gain a better understanding of the channels through which good governance in transparent firms create value, we follow the strategy of Giroud and Mueller (2011) to examine the relation between governance and investment activity. In particular, we use the same specification as in the operating performance regressions, except that the dependent variable is now either capital expenditures or some measure of acquisition activity. These additional estimation results are gathered in Table 2.14.

In columns [1] to [3], the dependent variable is capital expenditures (item #128) scaled by total assets (item #6). It is industry-adjusted by subtracting the industry median from the variable. Industry median is computed every year for each of the 48 FF industries. The dependent variable is trimmed at the 5th and 95th percentiles of its empirical distribution. All explanatory variables are lagged. The sample period is from 1990 to 2006. As is shown, the coefficient on G-index is large and significant for the transparent group, and it is small and insignificant for the opaque group. For example, in column [2], when transparency is measured by forecast dispersion, the coefficient is 0.044 (t-statistic = 2.36) for the transparent group, and it is 0.018 (t-statistic = 1.12) for the opaque group. This finding implies that good governance firms have less capital expenditures on average and this relation is more pronounced for transparent firms.

In columns [4] to [12], we investigate the relation between governance and acquisition activity measured by acquisition ratio, acquisition count, and acquisition likelihood. Acquisition ratio is the sum of the value of all acquisitions made by a firm in a given year divided by the firm's market capitalization in that year. Acquisition count is the number of acquisitions made by a firm in a given year. Acquisition likelihood is a dummy variable that equals one if the number of acquisitions made by a firm in a given year is non-zero and zero otherwise. These variables are constructed using data from the Securities Data Corporation's (SDC) database. As is displayed in the table, the coefficient on G-index is always large and significant for the transparent group, and it is smaller and insignificant or marginally significant for the opaque group. This evidence suggests that weak governance firms make more acquisitions, but this relation exists only for the transparent group. In other words, transparent firms with good governance make less acquisitions.

Considerable evidence in the literature shows the existence of negative announcement return and negative abnormal performance by acquiring firms. This evidence together with our investment activity test results suggest that good governance in transparent firms creates value by reducing agency costs. This provides a possible explanation for the pattern across transparency groups that we consistently find in our firm value, operating performance, and stock return tests.

2.7 Conclusion

Since Jensen and Meckling (1976), economists have devoted much effort to studying a firm's governance and a firm's information environment. Recent cases of poor governance as, e.g., in the scandals of Enron or Worldcom, have lead legislators to mandate rules that enforce more transparency, suggesting governance and transparency are substitutes. However, this is not clear theoretically, since there are compelling arguments for governance and transparency to be complements or substitutes. This paper aims at providing an empirical answer to the following question: Are firms with better governance (measured by a lower G-index) associated with better performance, on average, if they are also more transparent (e.g., measured by forecast dispersion)?

In essence, we document a complementary effect between transparency and governance by showing that the Democracy-Dictatorship hedge portfolio is associated with significantly positive abnormal returns when firms are transparent, while the hedge portfolio earns no abnormal return when firms are opaque. The positive effect of good corporate governance on stock returns is larger than that reported in GIM. For example, using the first principal component to combine the information of forecast error, forecast dispersion, and revision volatility, the hedge portfolio formed on this principal component earns a monthly abnormal return of 1.37% for value-weighted (1.28% for equal-weighted) portfolios, which is nearly twice as large as the alpha in the full sample. We also find that good governance firms, as measured by the G-index, have higher firm value and higher operating performance, but only among transparent firms. When we examine the channels through which good governance in transparent firms creates value, we find that good governance firms make less inefficient investments and less value-destroying acquisitions, but again mainly when they are also transparent firms.

Our findings suggest several broader issues. First, they suggest for corporate managers that firms thus far may have failed to recognize the complementarity of governance and transparency. Hence when firms improve their governance, they could also improve their disclosure policy to perhaps benefit more from improved governance. Second, academic researchers, who study the impact of corporate policies on equity returns, could benefit from recognizing a firm's information environment, because transparency could be also independently important for amplifying other corporate policies' effect on firm performance. Third, this study hints at improving transparency (by, e.g., the Regulation Fair Disclosure in 2000 or the Sarbanes-Oxley Act of 2002) might help policy makers to accomplish better firm performance, while focusing on governance alone might not accomplish this goal. Finally, the subtle interaction of governance and transparency that influences equity prices (or performance) that we document is not yet fully understood theoretically. These issues should also be fruitful for future research.

2.8 References

Amel-Zadeh, Amir, and Yuan Zhang, 2011, "The economic consequences of financial reporting quality for the market for corporate control: evidence from financial restatements," Working paper, Columbia University.

Armstrong, Christopher S., Karthik Balakrishnana, and Daniel Cohen, 2012, "Corporate governance and the information environment: evidence from state antitakeover laws," *Journal of Accounting and Economics* 53, 185–204.

Bebchuk, Lucian A., Alma Cohen, and Allen Ferrell, 2009, "What matters in corporate governance?" *Review of Financial Studies* 22, 783–827.

Bebchuk, Lucian A., Alma Cohen, and Charles Wang, 2011, "Learning and the disappearing association between governance and returns," *Journal of Financial Economics*, forthcoming.

Bond, Philip, Alex Edmans, and Itay Goldstein, 2012, "The real effects of financial markets," *Annual Reviews of Financial Economics*, forthcoming.

Botosan, Christine A., and Marelene A. Plumlee, 2002, "A re-examination of disclosure level and expected cost of equity capital," *Journal of Accounting Research* 40, 21–40.

Bushman, Robert, Josph D. Piotroski, and Abbie A. Smith, 2004, "What determines corporate transparency?" *Journal of Accounting Research* 42, 207–252.

Carhart, Mark M., 1997, "On persistence in mutual fund performance," *Journal of Finance* 52, 57–82.

Core, John E., Wayne R. Guay, and Tjomme O. Rusticus, 2006, "Does weak governance cause weak stock returns? An examination of firm operating performance and investors expectations," *Journal of Finance* 61, 655–687.

Cremers, K. J. Martijn, and Vinay B. Nair, 2005, "Governance mechanisms and equity prices," *Journal of Finance* 60, 2859–2894.

Cremers, K. J. Martijn, Vinay B. Nair, and Urs Peyer, 2008, "Takeover defenses and competition, *Journal of Empirical Legal Studies* 5, 791–818.

Cremers, K. J. Martijn, Vinay B. Nair, and Kose John, 2009, "Takeovers and the cross-section of returns," *Review of Financial Studies* 22, 1409–1445.

Dechow, Patricia M., and Ilia D. Dichev, 2002, "The quality of accruals and earnings: The role of accrual estimation errors," *The Accounting Review* 77 (Supplement), 35–59.

Demsetz, Harold, and Kenneth Lehn, 1985, "The structure of corporate ownership: causes and consequences, *Journal of Political Economy* 93, 1155–1177.

Doidge, Craig, Andrew Karolyi, Rene M. Stulz, 2007, "Why do countries matter so much for corporate governance?" *Journal of Financial Economics* 86, 1–39.

Duchin, Ran, John Matsusaka, and Oguzhan Ozbas, 2010, "When are outside directors effective?" *Journal of Financial Economics* 96, 195–214.

Easterwood, John C., and Stacey R. Nutt, 1999, "Inefficiency in analysts' earnings forecasts: systematic misreaction or systematic optimism?" *Journal of Finance* 54, 1777–1797.

Fama, Eugene F., and Kenneth R. French, 1997, "Industry costs of capital," *Journal of Financial Economics* 43, 153–193.

Ferreira, Miguel A., and Paul A. Laux, 2007, "Corporate governance, idiosyncratic risk, and information flow," *Journal of Finance* 62, 951–989.

Fishman, Michael J., and Kathleen M. Hagerty, 1989, "Disclosure decisions by firms and the competition for price efficiency," *Journal of Finance* 44, 633–646.

Francis, Jennifer, Ryan LaFond, Per Olsson, and Katherine Schipper, 2005, "The market pricing of accruals quality," *Journal of Accounting and Economics* 39, 295–327.

Giroud, Xavier, and Holger M. Mueller, 2011, "Corporate governance, product market competition, and equity prices," *Journal of Finance* 66, 563–600.

Givoly, Dan, and Josef Lakonishok, 1979, "The information content of financial analysts' forecasts of earnings," *Journal of Accounting Research* 1, 165–185.

Gompers, Paul A., Joy L. Ishii, and Andrew Metrick, 2001, "Corporate governance and equity prices," NBER working paper #8449.

Gompers, Paul A., Joy L. Ishii, and Andrew Metrick, 2003, "Corporate governance and equity prices," *Quarterly Journal of Economics* 118, 107–155.

Gu, Lifeng, 2012, "Transparency, takeover likelihood, and the cross-section of stock returns," Working paper, University of Illinois.

Hail, Luzi, and Christian Leuz, 2009, "Cost of capital effects and changes in growth expectations around U.S. cross-listings," *Journal of Financial Economics* 93, 428–454.

Hand, John R. M., 2003, "Profits, losses and the nonlinear pricing of internet stocks," in John R. M. Hand and Baruch Lev, eds.: *Intangible Assets: Values, Measures and Risks* (Oxford University Press, Oxford).

Hermalin, Benjamin E., and Michael S. Weisbach, 1988, "The determinants of board composition," *RAND Journal of Economics* 19, 589–606.

Hermalin, Benjamin E., and Michael S. Weisbach, 2007, "Transparency and corporate governance," NBER working paper #12875.

Hou, Kewei, and David T. Robinson, 2006, "Industry concentration and average stock returns," *Journal of Finance* 61, 1927–1956.

Jensen, Michael C., and William H. Meckling, 1976, "Theory of the firm: managerial behavior, agency costs and ownership structure," *Journal of Financial Economics* 3, 305–360.

Lang, Mark H., and Russel J. Lundholm, 1996, "Corporate disclosure policy and analyst behavior," *Accounting Review* 71, 467–492.

Lang, Mark H., Karl V. Lins, and Mark Maffett, 2010, "Transparency, liquidity, and valuation: international evidence," Working paper, University of Pennsylvania.

Larcker, David F., Richardson, Scott A., Tuna, Irem, 2007, "Corporate governance, accounting outcomes, and organizational performance," *Accounting Review* 82, 963–1008.

Leuz, Christian, and Robert E. Verrecchia, 2000, "The economic consequences of increased disclosure," *Journal of Accounting Research* 38, 91–124.

Leuz, Christian, and Felix Oberholzer, 2006, "Political relationships, global financing, and corporate transparency: evidence from Indonesia," *Journal of Financial Economics* 81, 411–439.

Lim, Terence, 2001, "Rationality and analysts' forecast bias," *Journal of Finance* 56, 369–385.

Marquardt, Carol A., and Emanuel Zur, 2011, "The role of accounting quality in the M&A market," Working paper, Baruch College.

Martin, Xiumin, and Ron Shalev, 2011, "Target firm-specific information and expected synergies in acquisitions, Working paper, Washington University in St. Louis.

McNichols, Maureen F., 2002, "Discussion of the quality of accruals and earnings: the role of accrual estimation errors," *The Accounting Review* 77 (Supplement), 61–69.

McNichols, Maureen F., and Stephen Stubben, 2011, "The effect of target-firm accounting quality on valuation in acquisitions," Working paper, Stanford University.

Mukherjee, Abhiroop, 2011, "Are control rights valuable when shareholders lack information?" Working paper, Hong Kong University of Science and Technology.

Novaes, Walter, and Luigi Zingales, 1995, "Capital structure choice when managers are in control: entrenchment versus efficiency," NBER working paper #5384.

Pastor, Lubos, and Robert F. Stambaugh, 2003,"Liquidity risk and expected stock returns," *Journal of Political Economy* 111,642–685.

Paul Jonathan M., 1992, "On the efficiency of stock-based compensation, *Review Financial Studies* 5, 471–502.

Shawn, Thomas, 2002, "Firm diversification and asymmetric information: evidence from analysts' forecasts and earnings announcements," *Journal of Financial Economics* 64, 373–396.

Shleifer, Andrei, and Robert Vishny,, 1997, "A survey of corporate governance,"*Journal of Finance* 52, 737–783.

Singh, Rajdeep, and Vijay Yerramilli, 2009, "Can transparency be too much of a good thing?," Working paper, University of Minnesota.

Stulz, Rene M.,1988, "Managerial control of voting rights: financing policies and the market for corporate control," *Journal of Financial Economics* 20, 25–54.

Teoh, Siew Hong, and T. J. Wong, 2002, "Why new issues and high-accrual firms underperform: the role of analysts' credulity," *Review of Financial Studies* 15, 869–900.

Zhang, X. Frank, 2006, "Information uncertainty and stock returns," *Journal of Finance* 61, 105–137.

Zwiebel, Jeffery, 1996, "Dynamic capital structure under managerial entrenchment," *American Economic Review* 86, 1197–1215.

2.9 Tables and Figures

Table 2.1: Summary Statistics

This table reports summary statistics on the empirical relation between the three transparency proxies and the G-index. The three transparency proxies are forecast dispersion, forecast error, and revision volatility. The definition of these variables is described in the data section. G-index is the governance index introduced in GIM and constructed using the 24 corporate provisions from the IRRC database. Firms are sorted into three sub portfolios (i.e., lowest tercile, medium tercile, and highest tercile) based on the distribution of their transparency proxies in both Democracy and Dictatorship portfolios. Then we compute the empirical distribution of transparency proxies between firms in the Democracy and Dictatorship portfolios that are in the same tercile. Panel A presents statistics for the sample period from 1990 to 2006. Panel B presents statistics for the sample period from 1990 to 1999.

Panel A: Distribution of transparency proxies in Democracy and Dictatorship portfolios (1990 to 2006)

		Democracy Portfolio			Dictatorship Portfolio		
		Mean	Median	Range	Mean	Median	Range
	Lowest Tercile	0.001	0.001	[0.0001, 0.002]	0.001	0.001	[0.0001, 0.002]
Forecast Dispersion	*Medium Tercile*	0.003	0.003	[0.002, 0.006]	0.003	0.003	[0.002, 0.004]
	Highest Tercile	0.01	0.01	[0.006, 0.04]	0.01	0.01	[0.004, 0.04]
	Lowest Tercile	0.0005	0.0005	[0.0001, 0.01]	0.0005	0.0006	[0.0001, 0.01]
Forecast Error	*Medium Tercile*	0.003	0.003	[0.001, 0.006]	0.002	0.002	[0.001, 0.004]
	Highest Tercile	0.02	0.01	[0.006, 0.05]	0.01	0.01	[0.004, 0.05]
	Lowest Tercile	0.001	0.001	[0.0001, 0.002]	0.001	0.001	[0.0001, 0.002]
Revision Volatility	*Medium Tercile*	0.004	0.003	[0.002, 0.006]	0.003	0.003	[0.002, 0.005]
	Highest Tercile	0.01	0.01	[0.006, 0.04]	0.01	0.01	[0.005, 0.04]

Panel B: Distribution of transparency proxies in Democracy and Dictatorship portfolios (1990 to 1999)

		Democracy Portfolio			Dictatorship Portfolio		
		Mean	Median	Range	Mean	Median	Range
	Lowest Tercile	0.001	0.001	[0.0001, 0.002]	0.001	0.001	[0.0001, 0.002]
Forecast Dispersion	*Medium Tercile*	0.003	0.003	[0.002, 0.005]	0.003	0.003	[0.002, 0.004]
	Highest Tercile	0.01	0.01	[0.005, 0.04]	0.01	0.01	[0.004, 0.04]
	Lowest Tercile	0.0004	0.0004	[0.0001, 0.001]	0.0005	0.0005	[0.0001, 0.001]
Forecast Error	*Medium Tercile*	0.003	0.003	[0.001, 0.005]	0.002	0.002	[0.001, 0.004]
	Highest Tercile	0.02	0.01	[0.005, 0.05]	0.01	0.01	[0.004, 0.05]
	Lowest Tercile	0.001	0.001	[0.0001, 0.002]	0.001	0.001	[0.0001, 0.002]
Revision Volatility	*Medium Tercile*	0.003	0.003	[0.001, 0.006]	0.003	0.003	[0.001, 0.005]
	Highest Tercile	0.01	0.01	[0.006, 0.04]	0.01	0.01	[0.005, 0.04]

Table 2.2 : Trading Strategies: Full Sample

This table reports the alphas for regressions of monthly excess returns to a hedge portfolio that takes a long position in Democracy firms and a short position in Dictatorship firms on an intercept (α), the market factor ($RMRF$), the size factor (SMB), the book-to-market factor (HML), and the momentum factor (UMD). $RMRF$, SMB, and HML factors are taken from Kenneth French's website. UMD factor is constructed according to Carhart (1997). G-index is the governance index introduced in GIM and constructed using the 24 corporate provisions from the IRRC database. Firms with a G-index of 5 or less are referred to as democratic firms and Firms with a G-index of 14 or more are referred to as dictatorship firms. Panel A represents the value-weighted results, for which the monthly returns are value-weighted by the market capitalization at the end of previous month. Panel B reports the equal-weighted results. In each panel, the original results from GIM, the replication results and the results using our sample are presented. The sample period ranges from September 1990 through December 1999. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Value-Weighted Democracy-Dictatorship Hedge Portfolios					
	α	*RMRF*	*SMB*	*HML*	*UMD*
GIM (2003)	0.71***	–0.04	–0.22	–0.55	–0.01
	(2.73)	(0.57)	(2.44)	(5.50)	(0.14)
Replication	0.67***	–0.04	–0.24	–0.54	0.02
	(2.67)	(0.57)	(2.71)	(5.27)	(0.26)
Our Sample	0.68**	–0.05	–0.24	–0.56	0.03
	(2.57)	(0.60)	(2.64)	(5.17)	(0.47)

Panel B: Equal-Weighted Democracy-Dictatorship Hedge Portfolios					
	α	*RMRF*	*SMB*	*HML*	*UMD*
GIM (2001)	0.45**	–0.00	0.23	–0.38	–0.16
	(2.06)	(0.01)	(3.02)	(4.30)	(2.79)
Replication	0.47**	–0.03	0.23	–0.38	–0.16
	(2.16)	(0.50)	(3.03)	(4.28)	(3.10)
Our Sample	0.35*	–0.01	0.19	–0.47	–0.11
	(1.77)	(0.19)	(2.73)	(5.80)	(2.27)

Table 2.3 : Trading Strategies: Transparency Proxies

This table reports abnormal returns for equal- and value-weighted Democracy-Dictatorship hedge portfolios using the Carhart four-factor model. Both Democracy and Dictatorship portfolios are divided into three terciles based on the three transparency proxies deflated by either lagged share price or lagged assets per share or absolute value of forecast mean: forecast dispersion, forecast error, and revision volatility. Then we form a Democracy-Dictatorship hedge portfolio for each transparency tercile every month and regress the monthly excess returns to each hedge portfolio on the market factor ($RMRF$), the size factor (SMB), the book-to-market factor (HML), and the momentum factor (UMD). The estimated intercept α is interpreted as the abnormal return of the trading strategy. Forecast error is defined as the absolute value of the difference between the actual annual earnings per share (EPS) and the mean of analyst forecasts. Forecast dispersion is defined as the forecast standard deviation across all analysts following the same firm in the same year. Revision volatility is computed as the standard deviation of the changes over the fiscal year in the median forecast from the preceding month. Panel A reports the α when transparency proxies are scaled by lagged share price. Panel B and Panel C show the results when transparency proxies are scaled by lagged assets per share and absolute value of forecast mean, respectively. The last two columns show the value-weighted abnormal returns to the Democracy (Long) and Dictatorship (Short) portfolios for the high transparency groups (i.e., lowest terciles). The sample period is from September 1990 to December 1999. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Equal-Weighted Hedge Portfolio			*Value-Weighted Hedge Portfolio*				
	Lowest Tercile	*Medium Tercile*	*Highest Tercile*	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*	*Long Leg*	*Short Leg*
	Panel A: Transparency Proxies Scaled by Lagged Price							
Forecast Dispersion	0.73**	0.40	0.05	0.99***	0.40	–0.21	0.81***	–0.18
	(2.41)	(1.31)	(0.18)	(3.00)	(0.87)	(0.50)	(3.92)	(0.79)
Forecast Error	0.80***	0.24	0.01	0.97***	0.09	0.06	0.76***	–0.21
	(2.88)	(0.86)	(0.01)	(2.99)	(0.22)	(0.15)	(3.88)	(0.81)
Revision Volatility	0.52**	0.27	0.13	0.73**	0.05	0.76	0.57***	–0.17
	(2.02)	(1.11)	(0.41)	(2.04)	(0.14)	(1.57)	(2.90)	(0.59)
	Panel B: Transparency Proxies Scaled by Lagged Assets							
Forecast Dispersion	0.61***	0.10	0.34	0.79**	0.36	0.29	0.50**	–0.29
	(2.59)	(0.34)	(0.65)	(2.32)	(0.81)	(0.40)	(2.41)	(1.19)
Forecast Error	0.56**	0.37	0.16	0.80**	0.87	0.08	0.48**	–0.31
	(2.24)	(1.22)	(0.36)	(2.31)	(1.85)	(0.13)	(2.43)	(1.26)
Revision Volatility	0.68***	0.20	0.14	0.59*	0.62	0.83	0.43*	–0.16
	(2.61)	(0.54)	(0.36)	(1.72)	(1.47)	(0.21)	(1.94)	(0.61)
	Panel C: Transparency Proxies Scaled by Forecast Mean							
Forecast Dispersion	0.51*	0.40	0.24	0.92***	0.46	0.13	0.66***	–0.17
	(1.75)	(1.39)	(0.74)	(2.67)	(1.15)	(0.24)	(3.22)	(0.82)
Forecast Error	0.44*	0.48	0.14	0.83**	0.67*	–0.04	0.45**	–0.30
	(1.70)	(1.56)	(0.44)	(2.55)	(1.93)	(0.08)	(2.15)	(1.16)
Revision Volatility	0.48*	0.24	0.43	0.74**	0.14	0.72	0.53***	–0.25
	(1.73)	(0.72)	(1.22)	(2.34)	(0.27)	(1.41)	(2.77)	(1.24)

Table 2.4 : Trading Strategies: Time-Series Average of Transparency Proxies

This table reports abnormal returns for equal- and value-weighted Democracy-Dictatorship hedge portfolios using the Carhart four-factor model. Both Democracy and Dictatorship portfolios are divided into three terciles based on the three transparency proxies' time-series averages: forecast dispersion, forecast error, and revision volatility. This leaves us with three Democracy-Dictatorship hedge portfolios every month. We then regress the monthly excess returns to each hedge portfolio on the market factor ($RMRF$), the size factor (SMB), the book-to-market factor (HML), and the momentum factor (UMD). The estimated intercept α is interpreted as the abnormal return of the trading strategy. Panel A reports the α when transparency proxies are scaled by lagged share price. Panel B and Panel C show the results when transparency proxies are scaled by lagged assets per share and absolute value of forecast mean, respectively. The sample period is from September 1990 to December 1999. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Equal-Weighted Hedge Portfolio			*Value-Weighted Hedge Portfolio*		
	Lowest Tercile	*Medium Tercile*	*Highest Tercile*	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*
	Panel A: Transparency Proxies Scaled by Lagged Price					
Forecast Dispersion	0.71**	0.41	0.03	0.86**	0.16	–0.16
	(2.44)	(1.41)	(0.11)	(2.47)	(0.42)	(0.39)
Forecast Error	0.75***	0.16	0.13	0.78**	0.26	0.18
	(2.90)	(0.50)	(0.42)	(2.30)	(0.65)	(0.44)
Revision Volatility	0.60**	0.53	–0.09	0.69*	0.09	0.70
	(2.09)	(1.67)	(0.28)	(1.86)	(0.27)	(1.35)
	Panel B: Transparency Proxies Scaled by Lagged Assets					
Forecast Dispersion	0.57**	0.56	0.23	0.76**	0.44	0.43
	(2.51)	(1.28)	(0.58)	(2.26)	(0.95)	(0.73)
Forecast Error	0.63***	0.41	0.28	0.75**	0.56	0.75
	(2.75)	(0.88)	(0.67)	(2.22)	(1.26)	(1.29)
Revision Volatility	0.67**	0.45	0.22	0.28	0.83*	0.46
	(2.13)	(1.38)	(0.64)	(0.65)	(1.92)	(0.92)
	Panel C: Transparency Proxies Scaled by Forecast Mean					
Forecast Dispersion	0.45*	0.35	0.24	0.83***	0.03	0.31
	(1.88)	(1.20)	(0.76)	(2.58)	(0.08)	(0.54)
Forecast Error	0.56**	0.30	0.18	0.74**	0.30	0.22
	(2.52)	(0.83)	(0.54)	(2.24)	(0.66)	(0.37)
Revision Volatility	0.59**	0.47	0.01	0.93***	0.24	0.18
	(2.01)	(1.67)	(0.03)	(2.60)	(0.49)	(0.31)

Table 2.5: Trading Strategies: Principal Component of Transparency Proxies

This table reports abnormal returns for equal- and value-weighted Democracy-Dictatorship hedge portfolios using the Carhart four-factor model. We first divide both Democracy and Dictatorship portfolios into three terciles based on the first principal component of forecast dispersion, forecast error, and revision volatility. This leaves us with three Democracy-Dictatorship hedge portfolios every month. We then regress the monthly excess returns to each hedge portfolio on the market factor ($RMRF$), the size factor (SMB), the book-to-market factor (HML), and the momentum factor (UMD). The estimated intercept α is interpreted as the abnormal return of the trading strategy. Panel A reports the α when transparency proxies are scaled by lagged share price. Panel B and Panel C show the results when transparency proxies are scaled by lagged assets per share and absolute value of forecast mean, respectively. The sample period is from September 1990 to December 1999. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Equal-Weighted Hedge Portfolio			*Value-Weighted Hedge Portfolio*		
	Lowest Tercile	*Medium Tercile*	*Highest Tercile*	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*
	Panel A: Transparency Proxies Scaled by Lagged Share Price					
α	1.28***	–0.22	0.04	1.37***	0.15	0.04
	(3.50)	(0.69)	(0.11)	(3.52)	(0.31)	(0.09)
	Panel B: Transparency Proxies Scaled by Lagged Assets					
α	0.76**	0.15	0.53	1.10**	0.88*	–0.20
	(2.47)	(0.41)	(0.96)	(2.45)	(1.88)	(0.29)
	Panel C: Transparency Proxies Scaled by Forecast Mean					
α	0.53*	0.28	0.35	0.89***	0.54	0.03
	(1.90)	(0.92)	(0.94)	(2.95)	(1.29)	(0.06)

Table 2.6 : Robustness (1)

This table reports abnormal returns for equal- and value-weighted Democracy-Dictatorship hedge portfolios using variants of regressions in Table 2.3. In Panel A, accruals quality is used to measure firm transparency. It is constructed as the standard deviation of the residuals of the following regression model:

$$\Delta WC_t = b_0 + b_1 CFO_{t-1} + b_2 CFO_t + b_3 CFO_{t+1} + b_4 \Delta Sales_t + b_5 PPE_t + \epsilon_t \ ,$$

where ΔWC_t is the change in working capital, CFO is cash from operations, $\Delta Sales_t$ is change in sales, and PPE is property, plant, and equipment. All variables are scaled by lagged assets. Each year this model is estimated for every firm using prior eight years of data. In Panel B, we use E-index instead of G-index as an alternative measure of corporate governance. E-index is developed in Bebchuk, Cohen and Ferrell (2009) and constructed based on 6 out of 24 corporate provisions. Firms with E-index value of 0 are assigned to "Democracy Portfolio" and firms with E-index value of 4 or more are assigned to "Dictatorship Portfolio". In Panel C, "new economy" firms are excluded firm the sample. "New economy" firms are classified in Hand (2003). In Panel D, we extend the sample period to 2011. In Panel E, results over the sample period of 2000 to 2011 are reported. The sample period is from September 1990 to December 1999. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Equal-Weighted Hedge Portfolio			*Value-Weighted Hedge Portfolio*		
Panel A: Accruals Quality	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*
Accruals Quality	0.82***	0.64	0.08	1.15***	0.70	–0.16
	(2.67)	(1.21)	(0.21)	(2.97)	(1.24)	(0.34)
Panel B: E-index						
Forecast Error	0.83***	0.34*	0.13	1.04***	0.56*	0.30
	(4.79)	(1.75)	(0.61)	(3.50)	(1.78)	(0.93)
Forecast Dispersion	0.59***	0.19	0.08	0.97***	0.01	0.57
	(3.00)	(0.95)	(0.33)	(3.21)	(0.03)	(1.44)
Revision Volatility	0.62***	0.12	0.23	0.79***	0.51**	0.42
	(3.42)	(0.64)	(1.14)	(2.96)	(2.03)	(1.02)
Panel C: Excluding "New Economy"						
Forecast Error	0.72***	0.21	–0.04	0.66**	0.07	0.02
	(2.65)	(0.75)	(0.12)	(2.00)	(0.16)	(0.04)
Forecast Dispersion	0.62**	0.40	–0.12	0.65**	0.36	–0.29
	(2.25)	(1.34)	(0.39)	(2.00)	(0.79)	(0.69)
Revision Volatility	0.58**	0.09	0.10	0.48	0.20	0.27
	(2.21)	(0.35)	(0.30)	(1.33)	(0.53)	(0.62)
Panel D: 1990-2011						
Forecast Error	0.50**	0.00	0.13	0.53**	–0.26	0.17
	(2.28)	(0.01)	(0.52)	(2.03)	0.73)	(037)
Forecast Dispersion	0.48**	–0.05	0.04	0.43	0.18	–0.24
	(2.23)	(0.23)	(0.14)	(1.61)	(0.43)	(0.64)
Revision Volatility	0.34	0.00	0.20	0.25	0.03	0.45
	(1.55)	(0.01)	(0.83)	(0.89)	(0.10)	(1.14)
Panel E: 2000-2011						
Forecast Error	0.30	0.08	–0.14	0.06	–0.43	–0.27
	(1.48)	(0.22)	(0.28)	(0.13)	(1.32)	(0.91)
Forecast Dispersion	0.31	–0.11	–0.23	–0.11	0.07	–0.32
	(1.56)	(1.66)	(1.14)	(0.67)	(0.10)	(0.45)
Revision Volatility	0.18	0.14	0.03	–0.21	0.06	–0.44
	(0.99)	(0.40)	(0.07)	(0.66)	(0.09)	(0.68)

Table 2.7 : Robustness (2)

This table reports the abnormal returns for equal- and value-weighted Democracy-Dictatorship hedge portfolios in various robustness checks. In Panel A, we restrict the sample to firms with high (above-median) institutional ownership. In Panel B, we restrict the sample to firms with low (below-median) institutional ownership. The institutional ownership is defined in Cremers and Nair (2005) as the percentage of shares held by the 18 largest public pension funds listed in their paper. In Panel C and D, the sample is split into firms in low and high competitive industries. The measure of industry competition is a sales-based Herfindahl-Hirschman index ("HHI"). It is computed as the sum of squared market shares for firms in each of the 48 FF industries. Market shares are computed using firms' sales data in Compustat. The sample period is from September 1990 to December 1999. *t*-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Equal-Weighted Hedge Portfolio			*Value-Weighted Hedge Portfolio*		
Panel A: High Inst. Ownership	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*
Forecast Error	1.08***	0.34	0.24	1.41***	1.28***	–0.12
	(3.11)	(0.95)	(0.59)	(3.67)	(2.91)	(0.19)
Forecast Dispersion	1.14***	0.25	0.39	1.54***	0.80	0.28
	(2.77)	(0.74)	(0.96)	(3.53)	(1.65)	(0.47)
Revision Volatility	0.95**	0.02	0.53	1.46***	0.27	0.91
	(2.45)	(0.06)	(1.47)	(3.18)	(0.62)	(1.59)
Panel B: Low Inst. Ownership						
Forecast Error	0.67	–0.03	–0.29	0.56	0.16	–0.42
	(1.50)	(0.07)	(0.61)	(1.14)	(0.31)	(0.86)
Forecast Dispersion	0.58*	0.22	–0.17	0.27	–0.16	–0.40
	(1.75)	(0.45)	(0.34)	(0.61)	(0.31)	(0.73)
Revision Volatility	0.28	0.35	–0.13	0.27	–0.10	–0.49
	(0.85)	(0.84)	(0.27)	(0.58)	(0.21)	(0.85)
Panel C: Low industry competition						
Forecast Error	0.86***	0.41	0.01	1.97***	0.24	0.39
	(2.66)	(1.10)	(0.02)	(3.94)	(0.40)	(0.53)
Forecast Dispersion	0.95**	0.06	0.39	1.97***	0.44	0.46
	(2.53)	(0.15)	(0.05)	(3.68)	(0.75)	(0.56)
Revision Volatility	0.79**	0.23	0.28	1.42***	0.35	0.96
	(2.07)	(0.55)	(0.49)	(2.73)	(0.58)	(1.46)
Panel D: High industry competition						
Forecast Error	0.74**	0.44	–0.01	0.69*	0.38	–0.20
	(2.38)	(1.41)	(0.02)	(1.89)	(0.98)	(0.47)
Forecast Dispersion	0.72**	0.55*	–0.09	0.62*	0.33	–0.31
	(2.40)	(1.87)	(0.27)	(1.70)	(0.73)	(0.82)
Revision Volatility	0.78**	0.05	0.33	0.67*	0.06	0.36
	(2.54)	(0.15)	(0.98)	(1.73)	(0.18)	(0.83)

Table 2.8: Robustness (3)

This table continues reporting the abnormal returns for various robustness checks. In Panel A, we restrict the sample to firms with low (below-median) leverage ratio. In Panel B, we restrict the sample to firms with high (above-median) leverage ratio. Leverage ratio is defined as the long term debt divided by firm assets value. In Panel C and D, the sample is divided into low (below-median) and high (above-median) asset groups. The sample period is from September 1990 to December 1999. *t*-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Equal-Weighted Hedge Portfolio			*Value-Weighted Hedge Portfolio*		
Panel A: Low leverage ratio firm	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*
Forecast Error	0.94**	0.20	0.20	1.67***	0.64	0.16
	(2.27)	(0.50)	(0.50)	(4.51)	(1.35)	(0.30)
Forecast Dispersion	0.75**	0.34	0.27	1.58***	0.39	0.40
	(1.98)	(0.65)	(0.68)	(3.53)	(0.86)	(0.70)
Revision Volatility	0.87**	–0.03	0.51	1.22**	0.26	062
	(2.48)	(0.06)	(1.17)	(2.55)	(0.56)	(1.15)
Panel B: High leverage ratio firm						
Forecast Error	1.03**	0.78*	–0.04	1.05**	0.99*	–0.02
	(2.39)	(1.81)	(0.10)	(2.41)	(1.87)	(0.04)
Forecast Dispersion	1.35***	0.31	0.11	1.13***	0.58	0.09
	(3.35)	(0.82)	(0.23)	(2.84)	(1.00)	(0.15)
Revision Volatility	0.97**	1.01	0.34	0.97**	0.65**	0.15
	(2.49)	(1.53)	(0.33)	(2.42)	(2.06)	(0.58)
Panel C: Small asset firm						
Forecast Error	1.44***	0.58	0.26	1.59***	0.83	0.04
	(2.63)	(1.42)	(0.63)	(2.60)	(1.50)	(0.06)
Forecast Dispersion	1.58***	0.57	0.14	1.79***	0.46	–0.03
	(2.77)	(1.54)	(0.33)	(2.96)	(0.82)	(0.05)
Revision Volatility	1.13***	0.65	0.51	1.74***	0.07*	0.59
	(2.62)	(1.35)	(1.13)	(2.76)	(0.13)	(1.05)
Panel D: Large asset firm Panel						
Forecast Error	0.47	0.21	0.11	0.92***	0.76	–0.11
	(1.57)	(0.54)	(0.27)	(2.58)	(1.53)	(0.19)
Forecast Dispersion	0.80**	0.07	–0.07	1.03**	0.16	0.49
	(2.41)	(0.18)	(0.17)	(2.44)	(0.31)	(1.02)
Revision Volatility	0.51	0.09	0.19	1.08**	0.02	1.00
	(1.43)	(0.25)	(0.52)	(2.48)	(0.06)	(2.00)

Table 2.9 : Industry Effects

This table reports the alphas of the trading stratedy using industry-adjusted returns, which are obtained by subtracting the median industry returns from the individual firm's return. The industry return is calculated in each of the 48 FF industries. In Panel A, the Fama-French three-factor model is used. In Panel B, the Carhart four-factor model is employed. The sample period is from September 1990 to December 1999. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Equal-Weighted Hedge Portfolio			*Value-Weighted Hedge Portfolio*		
Panel A: Fama-French Three-Factor model with industry-adjusted returns	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*
Forecast Error	0.77***	0.15**	–0.15	1.20***	–0.34	–0.52
	(3.00)	(2.14)	(1.21)	(3.33)	(0.88)	(1.37)
Forecast Dispersion	0.69***	0.26	–0.15	0.86***	0.01	–0.42
	(2.99)	(0.96)	(0.57)	(2.80)	(0.03)	(1.06)
Revision Volatility	0.62***	0.03	0.13	0.75*	–0.13	0.24
	(2.69)	(0.14)	(0.49)	(1.97)	(0.39)	(0.60)
Panel B: Carhart Four-Factor model with industry-adjusted returns	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*
Forecast Error	0.81***	0.19	–0.05	1.20***	0.36	–0.46
	(3.35)	(0.84)	(0.21)	(3.28)	(0.92)	(1.20)
Forecast Dispersion	0.70***	0.32	–0.05	0.85***	0.03	–0.34
	(3.01)	(1.23)	(0.18)	(2.74)	(0.06)	(0.88)
Revision Volatility	0.63***	0.09	0.25	0.73*	–0.09	0.31
	(2.73)	(0.35)	(0.97)	(1.88)	(0.28)	(0.77)

Table 2.10 : Fama-MacBeth Return Regressions

This table summarizes Fama-MacBeth estimates from monthly cross-sectional regressions of individual stock returns on an intercept, either the G-index or a Democracy dummy or the interaction terms between the Democracy dummy and transparency dummies along with a set of control variables. The estimated model is given in Equation (2.3) (see Section 2.5.3). The Democracy dummy equals one if the firm is in the Democracy portfolio. It equals zero if the firm is in the Dictatorship portfolio. Control variables include book-to-market ratio, gross return from month $t-3$ to month $t-2$, gross return from month $t-6$ to month $t-4$, gross return from month $t-12$ to month $t-7$, firm size, book leverage, stock price, sales growth over previous five years, trading volume of NYSE or Amex stocks, trading volume of NASDAQ stocks, a NASDAQ dummy, an S&P 500 dummy, dividend yield, institutional ownership, product market competition (measured by a sales-based Herfindahl-Hirschman index), and firm idiosyncratic volatility. The definition of these variables can be found in the Appendix in GIM. All explanatory variables are lagged. For Brevity the coefficients on the control variables are not reported. Column [1] uses the full sample and Columns [2] to [5] use the sample including only Democracy and Dictatorship firms. Column [1] reports the coefficient on G-index. Column [2] reports the coefficient on the Democracy dummy. Column [3] reports the coefficients on interaction terms when forecast dispersion is used for the transparency dummies. Column [4] reports the coefficients on interaction terms when forecast error is used for the transparency dummies. Column [5] reports the coefficients on interaction terms when revision volatility is used for the transparency dummies. The sample period is from September 1990 to December 1999. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	[1]	[2]	[3]	[4]	[5]
G-index	–0.03				
	(1.45)				
Democracy		0.34			
		(1.51)			
Democracy × Lowest			0.64**	0.76***	0.78***
			(2.43)	(2.66)	(2.80)
Democracy × Medium			0.27	–0.12	–0.12
			(1.10)	(0.40)	(0.41)
Democracy × Highest			–0.43	–0.29	0.26
			(1.25)	(0.82)	(0.82)
Lowest Tercile			0.62**	0.44	0.73**
			(2.12)	(1.45)	(2.33)
Medium Tercile			0.16	0.48*	0.31
			(0.64)	(1.88)	(1.15)
Number of observations	124,052	19,467	19,467	19,467	19,467
Number of months	112	112	112	112	112

Table 2.11: Alternative Asset Pricing Models

This table reports abnormal returns for equal- and value-weighted Democracy-Dictatorship hedge portfolios using alternative asset pricing models. In Panel A, the market (capital asset pricing model) model is employed to replace the Carhart four-factor model. In Panel B, the Fama-French four-factor model is used. The Fama-French momentum factor is downloaded from Kenneth French's website. In Panel C, we augment the Carhart four-factor model with the liquidity factor of Pastor and Stambaugh (2003) and estimate this five-factor model. In Panel D, we extend the Carhart four-factor model by the takeover factor of Cremers, Nair, and John (2009) and estimate this five-factor model. In all Panels, we use the principal component of the three transparency proxies as measure of transparency. The sample period ranges from September 1990 to December 1999. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Equal-Weighted Hedge Portfolio			*Value-Weighted Hedge Portfolio*		
	Lowest Tercile	*Medium Tercile*	*Highest Tercile*	*Lowest Tercile*	*Medium Tercile*	*Highest Tercile*
	Panel A: Market model					
α	0.53*	–0.09	0.10	0.91*	0.76	–0.44
	(1.68)	(0.26)	(0.17)	(1.95)	(1.67)	(0.64)
	Panel B: Fama-French four-factor model					
α	0.66**	-0.16	0.37	0.85**	0.65	–0.31
	(2.41)	(0.46)	(0.70)	(2.24)	(1.52)	(0.45)
	Panel C: Five factor model including the liquidity factor					
α	0.81***	0.07	0.58	1.11**	0.73	-0.10
	(2.66)	(0.19)	(1.05)	(2.56)	(1.62)	(0.15)
	Panel D: Five factor model including the takeover factor					
α	0.75**	0.24	0.65	0.92**	0.64	0.32
	(2.49)	(0.55)	(0.99)	(2.16)	(0.91)	(0.43)

Table 2.12: Governance, Transparency, and Firm Value

This table reports the results from panel regressions of industry-adjusted Tobin's Q on an intercept, year and industry fixed effects, the G-index (or the interaction terms between the G-index and transparency dummies), firm size, firm age, S&P 500 dummy, and Delaware dummy. The control variables also include transparency dummies to account for the direct effect of the transparency terciles. Tobin's Q is the market value of assets divided by the book value of assets. Industry-adjusted Tobin's Q is calculated by subtracting the industry median in a given FF industry and year. Firm size is the logarithm of firm assets. Column [1] reports the coefficient on G-index. Column [2] reports the coefficients on interaction terms when forecast error is used. Column [3] reports the coefficients on interaction terms when forecast dispersion is used. Column [4] reports the coefficients on interaction terms when revision volatility is used. The sample period is from 1990 to 2006. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	[1]	[2]	[3]	[4]
G-index	–0.035***			
	(2.75)			
G-index × Lowest		–0.095***	–0.110***	–0.104***
		(3.35)	(3.70)	(3.78)
G-index × Medium		–0.023***	–0.009	–0.012
		(2.60)	(1.31)	(1.40)
G-index × Highest		0.008	0.008	0.006
		(1.06)	(0.71)	(0.54)
Year-FE	Yes	Yes	Yes	Yes
Industry-FE	Yes	Yes	Yes	Yes
Number of Obs.	12,792	12,792	12,792	12,792
Adj. R-squared	0.07	0.14	0.18	0.18

Table 2.13 : Governance, Transparency, and Operating Performance

This table reports the results from panel regressions of industry-adjusted measures of operating performance on an intercept, year and industry fixed effects, the interaction terms between the G-index and transparency dummies, firm size, firm age, S&P 500 dummy, Delaware dummy, and the logarithm of the book-to-market ratio. The control variables also include transparency dummies to account for the direct effect of the transparency terciles. Firm performance measure is either return on assets (ROA), return on equity (ROE), or net profit margin (NPM). ROA is defined as net income divided by the book value of assets (item #6), ROE is defined as net income divided by the book value of common equity (item #60), and NPM is defined as net income divided by sales (item #12). All explanatory variables are lagged. All performance variables are industry-adjusted, which is calculated by subtracting the industry median from the variable. Industry median is computed every year for each of the 48 FF industries. All dependent variables are trimmed at the 5th and 95th percentiles of their empirical distribution. Columns [1] to [3] report the coefficients when the dependent variable is ROA. Columns [4] to [6] report the coefficients when the dependent variable is ROE, and Columns [7] to [9] report the coefficients when the dependent variable is NPM. In columns [1], [4], and [7], forecast error is used to measure transparency, in columns [2], [5], and [8], forecast dispersion is used to measure transparency, and in columns [3], [6], and [9] revision volatility is used to measure transparency. All coefficients are multiplied by 100. The sample period is from 1990 to 2006 and the standard errors are clustered at the industry level. *t*-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	ROA			NPM			ROE		
	[1] FE	[2] FD	[3] RV	[4] FE	[5] FD	[6] RV	[7] FE	[8] FD	[9] RV
G-index × Lowest	–0.182**	–0.219***	–0.226***	–0.393*	–0.411*	–0.397*	–0.262**	–0.347**	–0.369***
	(2.28)	(2.72)	(2.62)	(1.76)	(1.84)	(1.86)	(2.17)	(2.54)	(2.62)
G-index × Medium	–0.082	–0.038	–0.058	–0.013	0.034	–0.076	–0.137	–0.131	0.028
	(1.36)	(0.68)	(1.00)	(0.10)	(0.28)	(0.54)	(1.14)	1.64)	(0.17)
G-index × Highest	0.007	0.033	0.036	–0.17	–0.144	–0.104	0.136	0.279	0.092
	(0.13)	(0.49)	(0.55)	(1.61)	(1.21)	(0.76)	(1.06)	(1.31)	(0.78)
Year-FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Obs.	2055	2055	2055	2063	2063	2063	2109	2109	2109
Adj.R-squared	0.33	0.34	0.35	0.12	0.13	0.12	0.21	0.21	0.21

Table 2.14 : Capital Expenditure and Acquisition Activity

This table reports the results from panel regressions of capital expenditures or some measures of acquisition activity on an intercept, year and industry fixed effects, the interaction terms between the G-index and transparency dummies, firm size, firm age, S&P 500 dummy, Delaware dummy, and the logarithm of the book-to-market ratio. The control variables also include transparency dummies to account for the direct effect of the transparency terciles. All explanatory variables are lagged. All dependent variables are trimmed at the 5th and 95th percentiles of their empirical distribution. In columns [1] to [3], the dependent variable is Capex which is defined as capital expenditures scaled by total assets. Capex is industry-adjusted by subtracting the industry median in a given 48 FF industry and year. In columns [4] to [6], the dependent variable is acquisition ratio, which is the sum of the value of all acquisitions made by a firm in a given year divided by the firm's market capitalization. In columns [7] to [9], the dependent variable is acquisition count, which is the number of acquisitions made by a firm in a given year. In columns [10] to [12], the dependent variable is acquisition likelihood, which is a dummy variable that equals one if the number of acquisitions made by a firm in a given year is non-zero and zero otherwise. In columns [1], [4],[7], and [10[], forecast error is used to measure transparency, in columns [2], [5],[8], and [11], forecast dispersion is used to measure transparency, and in columns [3], [6], [9], and [12], revision volatility is used to measure transparency. All coefficients are multiplied by 100. The sample period is from 1990 to 2006 and the standard errors are clustered at the industry level. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Capex			Acquisition Ratio			Acquisition Count			Acquisition Likelihood		
	[1] FE	[2] FD	[3] RV	[4] FE	[5] FD	[6] RV	[7] FE	[8] FD	[9] RV	[10] FE	[11] FD	[12] RV
G-index ×	0.043**	0.044**	0.049***	1.16***	1.07***	1.25***	9.89***	9.83**	10.23***	2.72**	3.99**	3.45***
Lowest	(2.30)	(2.36)	(2.71)	(2.94)	(2.68)	(3.36)	(8.50)	(8.50)	(8.27)	(2.47)	(3.75)	(3.16)
G-index ×	0.029	0.028	0.026	1.07***	0.96***	0.86**	4.62***	4.45***	6.22***	2.32**	3.00***	3.04***
Medium	(1.66)	(1.56)	(1.44)	(2.71)	(2.61)	(2.13)	(3.60)	(3.89)	(5.50)	(2.19)	(2.88)	(2.91)
G-index ×	0.018	0.018	0.008	0.61	0.71	0.69	3.01*	2.52*	1.16	1.08	0.16	0.85
Highest	(1.14)	(1.12)	(0.55)	(1.63)	(1.62)	(1.59)	(1.88)	(1.90)	(0.93)	(1.03)	(0.16)	(0.80)
Year-FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regression type	OLS	OLS	OLS	Tobit	Tobit	Tobit	Poisson	Poisson	Poisson	Probit	Probit	Probit
Number of Obs.	10395	10395	10395	15221	15221	15221	15221	15221	15221	15221	15221	15221
Adj. R-squared	0.26	0.25	0.26	0.16	0.16	0.16	0.17	0.17	0.18	0.13	0.13	0.13

Chapter 3

Takeover Likelihood, Firm Transparency and the Cross-Section of Returns

3.1 Introduction

Takeovers are an important market activities and have attracted significant attention from both practitioners and financial economists. These events not only provide profitable opportunities for investors, but also create a great research platform for researchers. In the prior literature, people have been proposing models to predict the probability of the occurrence of takeover events. The most recent study in this research area is Cremers, Nair, and John (2009) who propose a logit model with several explanatory variables and find positive premium associated with takeover likelihood. McNichols and Stubben (2011) also show that acquirer returns around the acquisition announcement are higher when target firms have higher-quality accounting information because acquirer firm can bit more effectively and pay less to the target. If better information environment reduces the uncertainty of the target valuation and thus facilitates takeovers, we would predict that better firm information environment can be an additional dimension that can have important impacts on firm's takeover likelihood.

This paper proposes a new empirical model for predicting firms' takeover probability and studies the link between takeover exposure with stock returns. Specifically I augment the logit model proposed in Cremers, Nair, and John (2009) with an additional variable which measures firms' transparency level and find that transparency does add additional predicting power to firms' takeover likelihood. The results are consistent for all three measures of firms' transparency level. Importantly, the new logit model has a slightly better performance than the model proposed by Cremers, Nair, and John (2009) in various tests throughout the paper.

First, when the transparency variable is added to the model, the logit estimation coefficient on this proxy is negative and highly statistically significant without posing a significant impact on other variables' effects. What is more, comparing the logit regression using the old model and that using the new model, I find that the Pseudo-R^2 of the regression increases about 20%, providing supportive evidence that the new augmented model is a better fit of real takeover activity.

Firms' predicted takeover probability over the next year is then constructed based on the logit estimation coefficients, and I compare the time-series of the average predicted takeover probability among firms in the top takeover probability quintile with the time-series of the real takeover occurrence rate for the top quintile. The curve for the predicted takeover probability can capture the trace of the real takeover rate pretty well, better than the curve for the predicted takeover probability formed using the logit estimation results with the old model since the correlation between the time-series of the predicted takeover likelihood and the real takeover rate is higher when the results from the new model is employed.

Throughout the analysis, the main proxy for firm's transparency level is accrual quality which is constructed as the standard deviation of the residuals of the estimation model that regresses changes in working capital on cash flows, changes in sales, and property, plant, and equipment. Lower value of the standard deviation of the residuals means better accrual quality and more transparent information environment. In order to ensure the robustness of the results, two alternative proxies are also used in the logit estimation. They are analyst forecast error and analyst forecast dispersion which are formed based on analysts' earnings forecasts from the Institutional Brokers' Estimates System (I/B/E/S). In particular, forecast error is defined as the absolute value of the difference between the actual annual earnings per share (EPS) and the mean of analyst forecasts. Forecast dispersion is defined as the forecast standard deviation across all analysts following the same firm in the same year. Lower value of forecast error and forecast dispersion implies higher transparency level.

Following Cremers, Nair, and John (2009), I also investigate the link between the takeover likelihood and the stock returns using portfolio analysis and find that firms with higher takeover likelihood are generally associated with higher stock returns over the sample period of 1991 to 2009. According to the predicted takeover likelihood, firms are sorted into quintiles or deciles. The long-short portfolio that buys firms in the top takeover probability quintile and sells firms in the bottom quintile earns a monthly equal-weighted abnormal return of 88 basis points after adjusting for common risk factors such as market factor, size factor, value factor and Carhart (1997) momentum factor. This monthly abnormal return increases to 140 basis points for the decile sorted long-short portfolio. Interestingly, the long-short portfolio formed using the new model generates higher average return and abnormal return than the long-short portfolio constructed using the old model. For example, the mean return to the decile spread portfolio formed using the old model and the new model is 120 basis points and 140 basis points, respectively. Although the difference is not remarkable, this pattern is true for all cases including the equal-weighted return, the value-weighted return, the decile sorted portfolio and the quintile sorted portfolio. The results confirm the old claim that the takeover exposure is indeed associated with positive premium and convey the new message that the new model is able to catch firm's real takeover exposure slightly better.

The fact that the logit estimation and the return calculations are over the same time period would cause in-sample "look-ahead" bias of the results because the information might reflect the period after the realization of the return. In order to correct this bias, I re-estimate the logit model using 10-year rolling windows. That is, the takeover probability in a year is constructed based on the logit coefficients estimated using all the observations over the past 10 years. This way ensures that firm's information is incorporated into the market before the return calculation period. This out-of-sample estimation confirms the previous findings and generates a monthly equal-weighted abnormal return of 80 basis points for the quintile sorted long-short portfolio and 126 basis points for the decile sorted long-short portfolio. This indicates that the in-sample bias should not have significant impacts on the analysis in the paper.

Cremers, Nair, and John (2009) proposes a takeover factor and provides supportive evidence that the takeover factor is able to capture part of the variations in the cross-section of equity returns. In order to compare the pricing ability of the new takeover factor with the old takeover factor, I construct the new takeover factor as the monthly return spread between the top quintile takeover likelihood portfolio and the bottom quintile takeover likelihood portfolio. For comparison, I also construct the old takeover factor in a similar way, but with portfolio sorted using the takeover probability constructed based on the old logit estimation results. Both takeover factors are able to bring the abnormal returns to the Fama-French 25 size and book-to-market sorted portfolios to a lower level after accounting for only the market factor or all four common factors, showing good pricing ability of both factors. Interestingly, the new takeover factor performs slightly better since including the new takeover factor further reduces the abnormal returns of the 25 portfolios in terms of magnitude and statistical significance, although the improvement is not huge. And this fact is true for both equal-weighted and value-weighted size and book-to-market sorted portfolios.

In order to access the premium associated with takeover exposure quantitatively, I conduct additional test using 100 Fama-French size and book-to-market sorted portfolios and compute the takeover premium in two steps. First, the portfolio beta on a specific factor is obtained as the loading on a particular factor in a multivariate regression of the excess return of each of the 100 portfolios on risk factors. Then the premium associated with different factors are obtained as the coefficients from the multivariate regression of the mean excess return of each portfolio on all portfolio betas. Notably, the premium associated with the new takeover factor is slightly higher than the premium associated with the old takeover factor when both Carhart (1997) four-factor model and the simple capital asset pricing model (CAPM) are employed as the benchmark model. This indirectly show that the new logit model can capture the potential takeover likelihood slightly better.

This paper contributes to the prior literature in several ways. First, it adds new perspectives to the development of empirical models predicting the probability of the occurrence of takeover events since the new model performs better than the old model in various analysis throughout the paper. As far as I know, it is the first article that links firm's information environment directly to the measurement of its takeover likelihood. Second, this paper relates to the research area that studies acquisition and valuations because it provides supportive evidence for the hypothesis that transparency can facilitate takeovers, which can be implied by several prior articles that claim better information leads to more precise valuation of the target firm. That is, this article is an indirect proof of the hypothesis proposed by priors. Third, it also contributes to the corporate governance literature since a recent article Gu and Hackbarth (2012) find that transparent firms benefit more from good corporate governance and if transparent firms are generally associated with higher takeover likelihood, then the high level of takeover exposure of transparent firms could stimulate the positive power of good corporate governance and thus generates more beneficial effects.

The rest of the paper is organized as follows. Section 2 provides details of the data source, variable definitions and summary statistics. Section 3 provides the logit estimation results of takeover likelihood and investigates the relation between takeover probability and stock returns. In section 4, I construct the new takeover factor and use the 25 Fama-French size and book-to-market portfolios and 100 Fama-French size and book-to-market portfolios as base assets to test its cross-sectional pricing performance. Comparisons between the new model and the old model are done throughout the analysis in sections 3 and 4. Finally, section 5 concludes the paper.

3.2 Data Sources and Definition of Variables

Throughout the paper, three main data sources are involved: the Securities Data Corporation's (SDC) database, which provides information for merger and acquisition cases, the North America COMPUSTAT Annual Files, which contains firm-level accounting data, and the Center for Research in Security Prices (CRSP) database, from which monthly stock returns data are obtained. Following Cremers, Nair, and John (2009), I only consider all completed or 100% completed takeover deals from the SDC database in our analysis and both friendly and hostile deals are included.[1] After matching the SDC database with COMPUSTAT, I obtain a sample of 2,420 takeover targets with non-missing estimation variables if I use all completed deals over the time period of 1991 to 2009,[2] and the number reduces to 2,229 takeover targets if I use 100% completed takeovers only.[3]

The logit model I employ to estimate firm's takeover probability involves several independent variables in the right hand side of the equation. They are defined as follows: Q is the ratio of the market value of assets to the book value of assets and the market value of assets is computed as the total assets plus the market value of common stock minus the sum of the book value of common equity and differed taxes. PPE is property, plant and equipment scaled by total assets. $Cash$ is the ratio of cash and short-term investments to total assets. $Size$ is measured by the the natural logarithm of firm market capitalization. $Leverage$ is the book value of debt scaled by total assets. ROA is the return on assets. $Industry$ is the dummy variable which equals one if in the previous year there was at least one takeover event in the firm's industry which is defined based on Fama-French 48 industry classifications. $Block$ is also a dummy variable that equals one if there is at least one institutional owner whose ownership stake in firm's outstanding shares exceeds 5% and zero otherwise. This variable is constructed by using the quarterly institutional (13F) holdings data from Thompson/CDA Spectrum.

In most of the analysis, firm's transparency level is measured by accrual quality (or shortly, AQ). Following McNichols (2002),[4] accrual quality is constructed as the standard deviation of the residuals of the following estimation model:

$$\Delta WC_t = b_0 + b_1 CFO_{t-1} + b_2 CFO_t + b_3 CFO_{t+1} + b_4 \Delta Sales_t + b_5 PPE_t + \epsilon_t \ , \tag{3.1}$$

where ΔWC_t is the change in working capital from year $t-1$ to year t. Specifically, it is computed as the

[1] Since the chance of the completion of a hostile takeover is low, the number of hostile completed deals in the sample is very small. Dropping hostile deals from the sample does not affect the logit estimation results.

[2] The construction of the transparency variable involves estimation over a long time window, so the starting point of the sample is delayed to be 1991 to obtain non-missing transparency variable. Otherwise the sample period can start from 1981.

[3] 100% comppleted takeovers refer to the deals in which 100% of the target is acquired.

[4] See also Dechow and Dichev (2002), Francis et al. (2005), McNichols and Stubben (2011).

increase in accounts receivable plus the increase in inventory minus the increase in accounts payable and accrued liabilities minus the increase in taxes accrued plus the increase (decrease) in other assets or liabilities. CFO is operating cash flow, $\Delta Sales_t$ is change in sales from year $t-1$ to year t, and PPE is property, plant, and equipment. All variables are scaled by lagged total assets. Each year this model is estimated for every firm using data of the prior eight years and the standard deviation of the residual is defined as the accrual quality.[5] A larger standard deviation means lower accrual quality and lower accounting transparency.

Alternative measures of firms' transparency level are constructed by using analysts' earning forecasts from the Institutional Brokers' Estimates System (I/B/E/S). Based on the earnings forecast data, I construct two transparency proxies: forecast error and forecast dispersion. In particular, forecast error is defined as the absolute value of the difference between the actual annual earnings per share (EPS) and the mean of analyst forecasts. Forecast dispersion is defined as the forecast standard deviation across all analysts following the same firm in the same year. These two variables are all standard in the literature and are frequently used by researchers in accounting and finance.[6] To make these measures of transparency comparable across firms, I deflate them by lagged stock price.[7]To ensure the reliability of these measures, I require that there are at least three different analysts providing forecasts for the firm during the year. To limit the influence of coding errors and outliers on the results, I remove observations for which forecast error is larger than 10% of the share price at the beginning of the fiscal year (approximately 2% of the sample).[8]

Table 3.1 presents the summary statistics for the independent variables used in the logit estimation. Specifically, the mean values of those variables for both non-target and target firms over the period from 1991 to 2009 are provided to see how these two groups differ in terms of the mean of those variables. As is shown in the table, for the sample using all completed takeovers, almost all variables for the target group are significantly different from those for the non-target group except PPE, $Cash$ and ROA. For example, the mean of Q for the target group is 1.967, while the mean of Q for the non-target group is 2.329, and the difference of the mean is highly statistically significant with the t-statistic of 8.675. This is reasonable because the acquirors are more likely the ones that have high valuations and seek good investment opportunities. The mean of the variable *industry* for the target group is 0.945 and it is 0.904 for the non-target group, and the difference between them is also very significant with a t-statistic of 10.543. This is consistent with the fact that takeovers are likely to be industry clustered. The difference of the mean of another variable $BLOCK$ is also highly significant with the target group having a higher average level of institutional ownership. This is consistent with Cremers and Nair (2005) who claim that the existence of block holders can facilitate takeovers. Importantly, the mean difference of the variable of our interest *Opacity* is highly statistically significant with the target group having lower value than the non-target group.[9] Specifically, the mean for the target group is 0.032 and it is 0.045 for the non-target group. The t-statistic for the mean difference is 11.662. This is certainly consistent with McNichols and Stubben (2011) who propose that better

[5] I also estimate the model using data of the prior ten years or twelve years and the results do not change qualitatively.

[6] See, e.g., Givoly and Lakonishok (1979), Lang and Lundholm (1996), Thomas (2002), and Zhang (2006).

[7] I also deflate those variables by forecast mean or total assets and consistent results are obtained.

[8] See, e.g., Easterwood and Nutt (1999), Lim (2001), Teoh and Wong (2002), and Giroud and Mueller (2011).

[9] Because opacity is measured by accrual quality, which is the standard deviation of the residual of a regression, lower value of accrual quality means lower level of opacity or higher level of transparency.

information environment or higher transparency level can also facilitate takeovers because potential bidders can have a better evaluation of the target. In sum, the summary statistics in this table provide important information about what might be crucial determinants of the probability of the occurrence of a takeover event and these information will be reflected in the logit estimation results in the following section.

Other data sources are also used throughout the analysis. Monthly observations for the standard Fama-French risk factors are available from Kenneth French's website. The momentum factor is constructed according to the procedure in Carhart (1997). Monthly average returns to different test portfolios are also downloaded from Kenneth French's website. The details will be described in later sections.

3.3 Takeover Probability

In this section, I use the new logit model to estimate firms' takeover probability in the next year and study the link between the takeover exposure and firms' equity return by computing the returns to the quintile or decile portfolios constructed based on firm's predicted takeover likelihood which is formed using the coefficients from the logit estimation. Comparisons between the results from the new model and the results from the old model are tabulated to test the augmented impact of the new variable opacity on the prediction of takeover likelihood.

3.3.1 Logit Estimation

The summary statistics in Table 3.1 show that target firms have a number of characteristics that are different from non-target firms. However, we are not clear about how these variables combine together to predict takeover probabilities. Following Cremers, Nair, and John (2009) and others, I now estimate the probabilities of being taken over in the next period using a logit model. I assume that the marginal probability of becoming a target over the next period follows a logistic distribution and is given by the following equation

$$P(T_{it} = 1) = \frac{1}{1 + exp(-\alpha - \beta X_{it-1})} \tag{3.2}$$

where T_{it} is a dummy variable that equals one if the firm is a target in year t, and X_{it-1} is a vector of explanatory variables known at the end of the previous year. The elements of X_{it-1} include Q, PPE, $Cash$, $Size$, $Leverage$, ROA, $Industry$, $Block$ and $Opacity$. Detailed definition of those variables are provided in the data section. All those variables except $Opacity$ have been used by researchers in the prior literature to understand and predict takeover events.[10] I augment the model by adding $Opacity$ as an additional predicting variable since several articles in the literature show that firms' transparent environment can facilitate takeovers because acquiring firms can bid more effectively and the expected synergies are larger for target firms that are more transparent (see, e.g., Martin and Shalev (2011) or McNichols and Stubben (2011)). Thus transparent firms are more attractive targets and takeover attempts are more likely. All the Compustat variables are industry adjusted by subtracting the median value of the empirical distribution from the data. Year dummy variables are also included in the logit regression to account for the time fixed effect.[11]

[10]see, for example, Ambrose and Megginson (1992), Cremers, Nair, and John (2009).

[11]I also estimate the model without the year dummy variables and the results are similar.

The logit model is estimated once using all the observations from 1991 to 2009[12] and the regression coefficients are reported in Table 3.2. Model 1 refers to the logit model used in the prior literature and model 2 refers to the new model with all the variables included in model 1 and the additional predicting variable: *Opacity*. For both models, I report the estimation coefficients for each variable. As is displayed, when using model 1, variables Q (the market-to-book ratio), *Block* (more than 5% ownership stake dummy), *Industry* (the dummy variable to capture the clustering of takeover activity within industry), and *Size* have large t-statistics and thus are extremely statistically significant.[13] This shows that target firms tend to have low market-to-book ratio, high institutional ownership and small size and it is likely that these firms are in industries with takeover occurance in the previous year. These facts are all consistent with the findings in the prior literature. However, the coefficients on *ROA* and *Leverage* are significant but with positive sign. This is not expected intuitively because firms with low leverage seem to be easier to be taken over since the deal involves less issue with debt holders and firms with low return on assets seem to be easier to become target because the low return could be the results of bad management. However, this is consistent with the prior literature because others also show positive signs associated with these variables sometimes with different samples. Thus these two variables seem not to have persistant predicting power of takeover probability.[14]

Model 2 refers to the augmented model with all the variables included in Model 1 and an additional variable *Opacity*. Notably, using model 2, the coefficient on the variable of interest *Opacity* is negative and highly statistically significant. Specifically, the coefficient is -7.468 with a t-statistic of 10.21. This shows that transparent firms have higher predicted takeover probability than opaque firms, all else equal. What's more, adding this additional variable does not diminish the effects of other variables at all and the coefficients on the variables included in model 1 are still highly significant with similar magnitude. For example, the coefficient for the variable *Industry* is 0.618 (t-statistic = 5.57) in model 1 and it increases to 0.671 with a t-statistic of 6.05 in model 2. The coefficient for the variable *Block* is 0.567 (t-statistic = 11.34) in model 1 and it is 0.549 with a t-statistic of 10.98 in model 2. The coefficient on the variable Q slightly decreases and the coefficient on the variable *Size* increases. This might be because there is correlation between the opacity variable and these two variables. Generally speaking, adding the new variable does not have a significant impact on the effects of the variables in the original model. Thus the results in this table confirm my hypothesis that transparency can facilitate takeovers and hence it should be another dimension that can affect firms' probability of being taken over in the next time period.

In order to see how well the old and new models fit the takeover data, I also report the Pseudo-R^2 for each estimation at the bottom of Table 3.2. It is notable that augmenting the original model with the *Opacity* variable raises the level of Pseudo-R^2 to some extent.[15] Specifically, it increases from 3.41% in model 1 to 4.20% in model 2. The increase is about 20% of the original value, indicating that the augmented model should fit the data better and have additional predicting power of the takeover likelihood.

[12] Since the estimation of the transparency measure accrual quality has to be constructed using the rolling windows regression. Reliable values of this measure can only be formed as early as 1991.

[13] The unreported p-value of those coefficients are less than 0.0001.

[14] I also try to use market leverage to replace book leverage, this does flip the sign of leverage from positive to negative sometimes, but not all the times, thus the predicting power of leverage is not so persistent.

[15] Cremers, Nair, and John (2009) reports a Pseudo-R^2 of 1.39% using model 1 with a sample of all completed deals from 1981 to 2004.

To examine whether the results are robust using different measures of transparency, I also run model 2 in Table 3.2 using alternative transparency proxies which are constructed based on analysts' earnings forecasts. They are termed forecast error and forecast dispersion. Specifically, forecast error is defined as the absolute value of the difference between the actual annual earnings per share (EPS) and the mean of analyst forecasts. Forecast dispersion is defined as the forecast standard deviation across all analysts following the same firm in the same year. Intuitively, higher level of forecast error and forecast dispersion imply lower level of firm transparency because analysts can collect more information about the future performance of the firm if the firm is willing to release more of its true status. [16] The logit estimation results using model 2 over the sample period of 1991 to 2009 are presented in Table 3.3.

According to Table 3.3, the estimation results are consistent with the previous findings. That is, the coefficients on the transparency proxies are statistically significant for both proxies forecast error and forecast dispersion and adding the new variable does not diminish the effects of other variables at all. Specifically, the coefficient for forecast error is –0.378 and it is statistically significant at 1% level. The coefficient for forecast dispersion is –0.320 with a t-statistic of 2.62. Notably, The Pseudo-R^2 also increases from model 1 to model 2. For example, the Pseudo-R^2 is 3.41% for the logit estimation using model 1 and it increases to 4.17% when augmenting the model with the transparency variable forecast error. Similar pattern is shown in the case of another transparency measure forecast dispersion. The Pseudo-R^2 increases to 4.11% in model 2. Therefore, the test results in Table 3.3 provide another piece of supportive evidence that firms' transparent environment does have additional predicting power on firms' future takeover probability.

To further examine the predictability of the new augmented model, I compare the predicted takeover likelihood with the realized takeover event rate. Firms' predicted takeover likelihood is computed using equation (2) and the logit estimation coefficients from model 2 in Table 3.2. Each year firms are sorted into quintiles or deciles based on the value of the predicted takeover probability. The real takeover event rate is calculated within each quintile or decile every year as the number of takeovers deflated by the total number of firms in that quintile or decile group. Table 3.4 reports the mean value of the predicted takeover likelihood and the realized takeover event rates over the sample period of 1991 to 2009. For comparison, I also perform the same analysis using model 1 to compute the predicted takeover probability and the statistics are reported in the right part of each panel in Table 3.4.

Panel A shows the results for quintile portfolios and Panel B shows the results for decile portfolios. As is shown in the left part of Panel A, when using model 2, the realized takeover rate is increasing monotonically with the predicted takeover likelihood. Specifically, the real takeover rate goes from 0.0117 for quintile 1 to 0.0515 for quintile 5 and the predicted takeover likelihood goes from 0.0152 in quintile 1 to 0.0527 in quintile 5. The correlation between the average predicted takeover likelihood and the realized takeover rate is as high as 0.98. Similar pattern can be found in the left part of Panel B for the decile sorted portfolio. The real takeover rate increases monotonically from 0.0112 in decile 1 to 0.0575 in decile 10 and the predicted takeover probability goes from 0.0077 in decile 1 to 0.0529 in decile 10. The correlation between these two measures is 0.97. Thus the results show that there are actually more takeover activities among firms with higher predicted takeover likelihood and this is an indication of the reasonably good predicting ability of the new logit model.

[16] These variables are all standard in the literature and are frequently used by researchers in accounting and finance.

The right part of the table reports the results when calculating the takeover probability using the estimation coefficients from the old logit model (Model 1 in Table 3.2). You can see that the predicted takeover likelihood generally follows the trace of the realized takeover event rate from quintile 1 to quintile 5 or from decile 1 to decile 10. For example, the realized takeover rate is 0.0114 in decile 1 and it increases to 0.0469 in decile 10 and the corresponding predicted takeover probability is 0.0149 in decile 1 and 0.0521 in decile 10. However, the correlation of these two measures is 0.67 for the quintile sorted portfolios and it is 0.94 for the decile sorted portfolios. These correlation values are lower than the corresponding values using model 2 as the takeover predicting model. Therefore, the results presented in this table provide supportive evidence that the new augmented model can actually fit the real takeover data slightly better and firm transparency level adds another dimension to the takeover event prediction.

Table 3.4 shows how fit the model is with the predicted takeover likelihood averaged over the sample period. However, an investigation of the relation between the predicted takeover likelihood and the real takeover rate over time would provide more valuable information. Thus in Figure 1, I plot the time-series of the average predicted takeover probability and the real takeover rates [17] for the top decile group (The decile with the highest level of predicted takeover likelihood) over the sample period of 1991 to 2009.[18] As is shown in the figure, the real takeover activity shows an up-trend in the 90s and then it goes down in early years in the 20th century and goes up again until the start of the financial crisis.[19] Although the time-series of the predicted takeover probability is less volatile than that of the actual takeover activity, it generally follows its trace pretty well and the correlation between these two series is as high as 0.86. Thus the predicted takeover probability can capture a reasonably large part of the variations of the realized takeover activity over time.

Again for comparison, I also plot the graphs when computing the predicted takeover likelihood using the estimation coefficients from model 1. As is shown in part two of Figure 1, the predicted takeover likelihood curve generally follows the path of the realized takeover rates over time, but it does not capture the details of the real takeover activity as well as the time-series of the predicted takeover probability constructed using model 2. This is also reflected in the correlation between these two series. In this case, the correlation is 0.76, which is lower than the correlation of the two time-series in part one of Figure 1.

In sum, the above test results show that augmenting the old logit model with the additional transparency variable produces better fit of the real takeover activity data. The new model shows better performance than the old one since it produces slightly higher Pseudo-R^2 for the logit estimation and a larger correlation between the time-series of the average predicted takeover probability and the realized takeover rate.

3.3.2 Returns to Takeover Probability Portfolios

Several articles in the prior literature show that the variation in firm takeover exposure is related to market conditions and thus equity returns. Cremers, Nair, and John (2009) is the first article that studies the link between takeover likelihood and equity returns. In order to see if the takeover probability constructed using

[17]The real takeover rate here is computed among firms with full logit estimation information. So it is slightly different from the actual takeover rates.

[18]The graph for other decile groups are also plotted, but not displayed here. All undisplayed graphs show reasonably good predictability of the real takeover activity by the predicted takeover likelihood.

[19]Figure 1 in Cremers, Nair, and John (2009) also shows similar trend in terms of the average value of deals every year from 1981 to 2004.

the new model has similar predictions for stock returns, I follow Cremers, Nair, and John (2009) to examine the relation between firms' takeover probability and stock returns using portfolio sorting approach.

The logit estimation coefficients from model 2 in Table 3.2 are used to compute the probability of being taken over in the next year by using equation (2). Then firms are sorted into quintile or decile portfolios every year according to the rank of their takeover likelihood. Monthly equal-weighted quintile portfolio returns as well as equal-weighted and value-weighted return to the long-short portfolio that holds a long position in firms with high takeover probability and a short position in firms with low takeover probability are reported in Table 3.5. For comparison, I also report the returns to the long-short portfolio constructed based on the takeover probability using model 1 at the bottom in each panel.

To investigate whether the portfolio returns can be captured by existing standard risk factors such as the market factor, the size factor, the value factor and the momentum factor, I also use Carhart (1997) four-factor model to adjust for different risk styles of the takeover probability-sorted portfolios and the abnormal portfolio returns alpha together with the statistical significance are also reported in Table 3.5. If the takeover-probability sorted portfolios are just the reflection of different combinations of the loadings on those existing factors, we would not expect any significant abnormal returns. However, if the portfolio returns can not be adjusted by existing four-factors, this implies an additional pricing factor.

The logit estimation in Table 3.2 uses all the observations over the period of 1991 to 2009 to compute the variable coefficients and the return calculations in this section are also over the same time period. This would cause in-sample "look-ahead" bias of the results because the information might reflect the period after the realization of the return. In order to correct this bias, I re-estimate the logit model using 10-year rolling windows.[20] For example, the takeover probability in year 2001 would be calculated using the logit estimation coefficients using all 10-year observations from 1991 to 2000. However, this out-of-sample estimation also has limitations. For example, the data requirements certainly shorten the portfolio return calculation periods by 10 years, which could also cause potential bias. For comparison purposes, the results using the 10-year rolling window estimation are also tabulated in the right part of Table 3.5.

Consistent with Cremers, Nair, and John (2009), I also find that the abnormal return to the takeover probability sorted portfolio generally increases as the takeover likelihood increases after I include firm transparency as an additional predicting variable in the logit estimation. As is shown, both mean return and abnormal return generally increase from Quintile 1 to Quintile 5, and the hedge portfolio that buys stocks with high takeover likelihood and sells stocks with low takeover likelihood earns an equal-weighted abnormal return of 88 basis points per month and it is highly statistically significant (t-statistic = 5.14). Not surprisingly, the equal-weighted abnormal return to the long-short decile portfolio is higher with 140 basis points per month and a t-statistic of 5.25.

The results using the takeover probability computed from the 10-year rolling window logit estimation are also very similar. In this case, the information used to construct the takeover likelihood is prior to the return period and the "look-ahead " bias is corrected. The monthly abnormal return α to the long-short quintile portfolio is 0.80% with a t-statistic of 3.22 and it is 1.26% with a t-statistic of 3.52 using the decile sorting

[20]In order to insure enough target observations in the logit estimation, we follow Cremers, Nair, and John (2009) to choose the 10-year rolling window for the logit estimation. Too short window will result in unstable and unreliable logit estimation results and too long window will leave me with only several years for the takeover probability calculation.

method. Thus this out-of-sample test also confirms the positive relation between firms' takeover likelihood and stock returns. The value-weighted alpha for the long-short portfolio are also computed and reported, but the results are a bit weaker in terms of statistical significance maybe because the time period is short in this case and it affects statistical significance.[21]

The returns to the long-short portfolios when the old model (model 1) is used to construct the takeover probability are also reported at the bottom of Table 3.5 for comparison. As is shown, the return spread and the four-factor alphas here are smaller than the return spread and the four-factor alphas when the takeover probability is constructed using the new model (model 2). For instance, the equal-weighted return to the long-short quintile portfolio is 89 basis points per month, which is 7 basis points lower than the return spread when model 2 is used. In the value-weighted case, the return spread is 0.84% and 0.67% when using model 2 and model 1, respectively. The equal-weighted four-factor alpha is 0.88% for model 2 and it is 0.79% for model 1. The value-weighted four-factor alpha is 0.42% and 0.21%, respectively.

In the case of decile sorting, the equal-weighted return to the long-short portfolio is 140 basis points per month, while it is 147 basis points per month when the new model is employed to construct the takeover probability. The difference of the value-weighted return spread is larger and it is 25 basis points. The equal-weighted four-factor alpha is 1.40% and 1.27% for model 2 and model 1, respectively. The value-weighted four-factor alpha is 0.68% and 1.55% for model 2 and model 1, respectively. In sum, the comparison within the table indicates that if takeover exposure is priced, then the augmented model can capture the real takeover likelihood slightly better than the old one and thus has a better prediction of the premium associated with takeovers.

In order to track the performance of the takeover probability sorted portfolio over time and compare this between model 1 and model 2, I compute the cumulative return to the long-short portfolio each year. Figure 2 illustrates the performance of the long-short portfolio that buys firms with top level of takeover likelihood and sells firms with bottom level of takeover likelihood over the sample period of 1991 to 2009. Firms are sorted into either quintiles or deciles and the time-series of the cumulative monthly return of the long-short portfolios are plotted and referred to as "LS1090" and "LS2080" in the graph. As is shown, the performance of all long-short portfolios is pretty consistent over time. In most of the years in the 90s and the 2000s, there are positive return spreads, with the most negative returns concentrated in a few years in late 2000s. Furthermore, the curves for the long-short portfolios when the takeover probability is formed based on the logit estimation results using model 1 are always below the corresponding curves in the case of model 2, indicating that the transparency variable in model 2 does have predicting power for takeover probability and including this in the logit model provides a slightly better catch of the takeover premium.

In sum, the results in this section are consistent with Cremers, Nair, and John (2009) who also show the positive relation between the likelihood of being taken over in the next period and the average stock returns. In this paper, I augment their model with an additional important variable that shows significant predicting power for takeovers and the long-short portfolio formed based on the logit estimation results using the new model generates a larger return spread than the long-short portfolio constructed based on the results using the old model. This implies that the premium associated with the takeover exposure can be larger if the

[21] Cremers, Nair, and John (2009) also find weaker value-weighted results.

takeover exposure can be better captured.

The significant abnormal returns to the long-short portfolios after accounting for common risk factors is a clear indication that there might be another factor that can also explain part of the variations in the cross-section of stock returns. Cremers, Nair, and John (2009) shows that the takeover factor created based on the old logit model has pricing power for certain base assets. Since the return spread using the new model is larger than that using the old model from Cremers, Nair, and John (2009), it would be interesting to see whether the takeover factor constructed based on the augmented logit model has improved performance in terms of pricing the cross-section of stock returns. Thus, in the next section, I recreate the takeover factor following the procedure in Cremers, Nair, and John (2009)and test its performance using 25 Fama-French size and book-to-market sorted portfolios as base assets. I also compute the premium associated with both the old takeover factor and the new takeover factor using 100 Fama-French size and book-to-market sorted portfolios.

3.4 Takeover factor

3.4.1 Construction of the new takeover factor

The takeover factor is constructed as the monthly equal-weighted portfolio return to the long-short portfolio that is long in firms with the top takeover likelihood and short in firms with the bottom takeover likelihood. Table 3.6 presents the summary statistics of the constructed takeover factor, Fama-French three factors (i.e., MKT, SML, HML), and the Carhart (1997) momentum factor (UMD).

Panel A in the table lists some basic statistics of the above five factors. The average monthly return of the takeover factor from 1991 to 2009 is 0.96% (t–statistic = 4.81). This confirms the previous results that there is a significant premium associated with takeovers over the sample period. Two other facts can be noticed from Panel A as well. First, mean of the TOP factor is higher than other factors over my sample period. Second, TOP factor is almost as volatile as market, size, book-to-market, and momentum factors. Panel B lists the correlation matrix of these factors. The takeover factor is negatively correlated with the market factor and the momentum factor. The correlation is –0.09 and –0.26, respectively. It also has a positive correlation of 0.32 with the size factor. This is consistant with the logit estimation results that smaller size firms tend to have higher takeover likelihood. Intuitively, it requires less resource for the potential bidder to take over a small size firm. The takeover factor is also positively correlated with the value factor and the correlation is 0.32.

3.4.2 Pricing 25 Fama-French Size and B/M portfolios

Next I turn to test the performance of the new takeover factor using 25 Fama-French size and book-to-market portfolios as the base assets.[22] Specifically, I estimate the following asset-pricing models:

$$R_t - R_f = \alpha + \beta_1 \times RMRF_t + \epsilon_t \tag{3.3}$$

$$R_t - R_f = \alpha + \beta_1 \times RMRF_t + \gamma \times TOP_t + \epsilon_t \tag{3.4}$$

$$R_t - R_f = \alpha + \beta_1 \times RMRF_t + \beta_2 \times SMB_t + \beta_3 \times HML_t + \beta_4 \times UMD_t + \epsilon_t \tag{3.5}$$

[22]The data set for the 25 portfolio returns are downloaded from Kenneth French's website.

$$R_t - R_f = \alpha + \beta_1 \times RMRF_t + \beta_2 \times SMB_t + \beta_3 \times HML_t + \beta_4 \times UMD_t + \gamma \times TOP_t + \epsilon_t \qquad (3.6)$$

Where $R_t - R_f$ is the excess return of the portfolio. I compare the regression intercept α and its t–statistic using the above four models. Since α represents the estimate of the expected excess returns unexplained by the risk factors in the asset pricing model, any amount of decrease in the magnitude or the significance of α can be an indicator of the improvement of the performance of the model.

Table 3.7 reports the mean excess return and the abnormal returns of the 25 Fama-French size and Book-to-Market equal-weighted portfolios adjusting for risk factors using different asset pricing models. The equal-weighted monthly mean excess return and the corresponding t-statistics are shown in Panel A as the benchmark to test the performance of different models. In Panel B and Panel D, where $CAPM$ model or four-factor model is used, the excess return is reduced to some extent. Comparing Panel B with Panel C, Panel D with Panel E, we can see that augmenting the traditional market model or four-factor model with the takeover factor further brings the CAPM alpha or four-factor alpha down in terms of magnitude and statistical significance. For example, for the smallest size and highest book-to-market portfolio, the monthly alpha goes from 0.89% to 0.31% and the t-statistic goes from 5.06 to 2.03 when the new takeover factor is included in the four-factor model. This indicates that the takeover factor has its own pricing power independent of the common risk factors. The test results using the value-weighted size and B/M portfolios are presented in Table 3.8 and similar conclusions can be drawn from the results there.

In order to investigate whether the new takeover factor constructed using the new model performs better than the original takeover factor formed based on the old model, I also construct the old takeover factor over my sample period using the estimation results of Model 1 in Table 3.2. An alternative five-factor model using this old takeover factor is estimated to compare with the pricing performance of the five-factor model employing the new takeover factor. The results using the equal-weighted size and book-to-market portfolios are presented in Table 3.9.

As the benchmark return for comparison, I report the mean excess return and the Carhart (1997) four-factor alpha in Panel A and Panel B, respectively. Panel C shows the abnormal returns adjusting for the five factors including the market factor, size factor, value factor, momentum factor, and the new takeover factor. Panel D represents the abnormal returns of the five-factor model including the old takeover factor. The new takeover factor and the old takeover factor are termed as "TOP" and "TOPO", respectively. Comparing the abnormal return alphas in Panel C with those in Panel D, we can see that the corresponding alphas from both panels are of comparative magnitude, with the model in Panel C having slightly better performance for some portfolios. For instance, for the smallest size and highest book-to-market portfolio, the monthly four-factor alpha is 0.89% (t-statistic = 5.06), the five-factor model including the new takeover factor is able to bring this alpha down to 0.31% (t-statistic = 2.03), while the five-factor model with the old takeover factor bring the alpha to 0.39% with a higher significance level. In this case, the t-statistic is 2.67, which is significant at 1% level, while using the new takeover factor, the alpha is significant at 5% level. Although for many other portfolios, the differences between the abnormal returns and the t-statistics in Panel C and Panel D are small, generally the five-factor model using the new takeover factor produces smaller or equal value of alpha and smaller or equal value of t-statistics. Thus, the new takeover factor does have slightly better pricing performance for the equal-weighted Fama-French size and book-to-market sorted portfolios than the old takeover factor.

The test results using the value-weighted Fama-French size and book-to-market portfolios are presented in Table 3.10. Similar insights can be obtained from this table. That is, comparing the alphas in Panel C with the alphas in Panel D, we can see that the five-factor model with the new takeover factor is able to bring the four-factor alpha to a lower level in some cases than the five-factor model with the old takeover factor, although the magnitude of the difference is small. For instance, for the smallest size and highest book-to-market portfolio, the five-factor alpha is 0.08% using the old takeover factor, and it is 0.02% using the new takeover factor. The corresponding t-statistics are 1.20 and 0.16, respectively. Although the difference is not huge, the performance of the new takeover factor is still slightly better. Thus, the results from these two tests show that the new takeover factor can capture the variations in the cross-sectional returns to a slightly better extent, indicating that the transparency variable does add extra predicting power to the logit model for the takeover likelihood.[23]

3.4.3 Premium associated with the takeover exposure

The results in previous tables show that the return spread of the long-short portfolio is larger when the new logit model is used in the logit estimation. This implies a larger premium associated with takeover exposure. In order to access this premium quantitatively, I use the 100 Fama-French size and book-to-market portfolio as base assets for this test.

The procedure of this test goes with two steps. First, the beta of each of the 100 portfolios is obtained by regressing the excess return of each of the 100 portfolios on risk factors and the loading or regression coefficient on the factor is considered as the beta of the portfolio associated with that particular factor. Then the premium associated with each factor is obtained as the coefficient of the regression of the mean excess return of those 100 portfolios on all portfolio betas. I compute both betas for the new takeover factor from the new augmented model and for the old takeover factor from the old model in Cremers, Nair, and John (2009) and compare the premiums associated with them. The results are presented in Table 3.11.

Panel A displays the coefficients using Carhart (1997) four-factor model as the benchmark model. As is shown in the first two columns in Panel A, all Fama-French three factors are priced, which is consistent with the findings in prior literature. When the takeover factor constructed using the old logit model is added to the regression, the coefficients on the other four factors only have slight change and there is a significant premium associated with the takeover factor over the sample period. Specifically, the annualized coefficient on the old takeover factor is 0.06 with a t-statistic of 4.23. This means that the premium for takeover exposure can be as high as 6% annually.[24] Notabally, while the new takeover factor is used to replace the old one in the third regression, this premium increases to 7% (t-statistic = 3.93) annually.

As a robustness check, I also perform similar tests using the simple capital asset pricing model (CAPM) as the benchmark model[25] and the results are presented in Panel B of Table 3.11. As is shown, the coefficient

[23]Considering that it is usually very difficult to have additional pricing power with a factor other than the existing known three factors, the slightly better improvement of the new takeover factor is already pretty good.

[24]Cremers, Nair, and John (2009) also reports 8% anual prerium associated with the takover factor constructed using the same logit model over the period of 1981 to 2004.

[25]As is shown in one of the prior tables, there is high level of correlations between the takeover factor and other common factors. Thus using the simple CAPM model to perform the tests again ensures that the results in Panel A are not driven by the correlations between those factors.

on the old takeover factor is 0.05 and significant (t-statistic = 2.55). while with the new takeover factor the premium goes up to 8% with a t-statistic of 4.05. This provides evidence consistent with the test results in Panel A.

I also report the R^2 of each regression at the bottom of each panel. As is shown in Panel A, the R^2 increases dramatically when the old takeover factor is included into the regression. Specifically, it goes from 21% to 41%. When the new takover factor is used as a replacement, the R^2 increases from 41% to 50%, which shows an improvement, although not significant. Similar pattern of R^2 can be found in Panel B. It jumps from 10% to 15% when the old takeover factor is employed and increases to 21% when the new takeover factor is used.

In sum, the test results in this section provide supportive evidence for the fact that takeover exposure is indeed associated with positive and significant premium in a quantitative way, further confirming the pricing ability of the takeover factor. The new takeover factor constructed using the model augmented with transparency variable is associated with a slightly higher premium, implying that the new logit model can capture the potential takeover likelihood slightly better.

3.5 Conclusion

Takeover events which are important parts of market activities have been drawing consistent attention from practitioners and financial researchers. These events provide profitable opportunities for both target and acquirer shareholders and other outside investors. Thus predicting the probability of the occurrence of these events is very interesting and people have been proposing different models on this in the prior literature. A recent article by Cremers, Nair, and John (2009) propose a model and show that the takeover factor they construct based on the predicted takeover probability is associated with positive premium. Gu and Hackbarth (2012) show that firm's information environment matters for the governance mechanisms to affect firm performance and propose that a firm's information environment might be an important additional dimension for gauging a firm's takeover likelihood.

This paper studies firms' takeover probability and its link with equity returns. I augment the logit model in Cremers, Nair, and John (2009) with an additional variable which measures firms' transparency level and find that transparency does provide additional predicting power of firm's takeover likelihood since the logit estimation coefficient on the transparency proxy has an expected sign and is statistically significant. The new augmented model not only produces consistent results as prior research, but also increases the Pseudo-R^2 of the logit regression by 20%, indicating that the new model fits the real takeover activity better. This finding is robust when using alternative two measures of transparency constructed based on analysts earnings forecasts: forecast error and forecast dispersion.

Following Cremers, Nair, and John (2009), I investigate the relation between takeover probability and stock returns and also find that higher takeover likelihood is generally associated with higher returns after adjusting for common risk factors such as the market factor, the size factor, the value factor, and the Carhart (1997) momentum factor. Interestingly, the long-short portfolio constructed based on the takeover probability formed using the logit estimation results using the new model generates higher mean return and abnormal returns than the long-short portfolio formed using the estimation results of the old model.

The new takeover factor constructed as the return to the long-short portfolio that buys stocks in the top takeover likelihood quintile and sells stocks in the bottom takeover likelihood quintile is able to reduce the abnormal return to the 25 Fama-French size and book-to-market portfolio further to a lower level than the takeover factor constructed based on the old model. What is more, the premium associated with the new takeover factor is slightly higher than the premium associated with the old takeover factor. These results indicate that the new takeover factor has slightly better pricing ability in the cross-section of stock returns.

This paper contributes to the literature in several ways. First, it adds new perspectives directly to the takeover likelihood prediction models by considering firm's information environment which is important but yet overlooked by priors and thus provides a better prediction of the real takeover events. Second, it provides indirect evidence for the claims proposed in several prior articles that acquirers can have more precise valuation of the target and thus bid more effectively if the target firm has better information quality. Third, this paper relates to the corporate governance literature since it provides supportive evidence for a recent article Gu and Hackbarth (2012) who find that transparent firms benefit more from good corporate governance. If transparency can facilitate takeovers, then the high level of takeover exposure of transparent firms could stimulate the positive effects of good corporate governance on firm performance and therefore transparent firms can benefit more from good corporate governance.

3.6 References

Armstrong, Christopher S., Karthik Balakrishnana, and Daniel Cohen, 2011, "Corporate governance and the information environment: evidence from state antitakeover laws," *Journal of Accounting and Economics*, Forthcoming.

Bebchuk, Lucian A., Alma Cohen, and Allen Ferrell, 2009, "What matters in corporate governance?" *Review of Financial Studies* 22, 783–827.

Botosan, Christine A., and Marelene A. Plumlee, 2002, "A re-examination of disclosure level and expected cost of equity capital," *Journal of Accounting Research* 40, 21–40.

Bushman, Robert, Josph D. Piotroski, and Abbie A. Smith, 2004, "What determines corporate transparency?" *Journal of Accounting Research* 42, 207–252.

Carhart, Mark M., 1997, "On persistence in mutual fund performance," *Journal of Finance* 52, 57–82.

Cremers, K. J. Martijn, and Vinay B. Nair, 2005, "Governance mechanisms and equity prices," *Journal of Finance* 60, 2859–2894.

Cremers, K. J. Martijn, Vinay B. Nair, and Kose John, 2009, "Takeovers and the cross-section of returns," *Review of Financial Studies* 22, 1409–1445.

Dechow, Patricia M., and Ilia D. Dichev, 2002, "The quality of accruals and earnings: The role of accrual estimation errors," *The Accounting Review* 77 (Supplement), 35–59.

Easterwood, John C., and Stacey R. Nutt, 1999, "Inefficiency in analysts' earnings forecasts: systematic misreaction or systematic optimism?" *Journal of Finance* 54, 1777–1797.

Fama, Eugene F., and Kenneth R. French, 1992, "The cross-section of expected stock returns," *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, "Common risk factors in the returns of bonds and stocks," *Journal of Financial Economics* 33, 3–56.

Giroud, Xavier, and Holger M. Mueller, 2011, "Corporate governance, product market competition, and equity prices," *Journal of Finance* 66, 563–600.

Givoly, Dan, and Josef Lakonishok, 1979, "The information content of financial analysts' forecasts of earnings," *Journal of Accounting Research* 1, 165–185.

Gompers, Paul A., Joy L. Ishii, and Andrew Metrick, 2003, "Corporate governance and equity prices," *Quarterly Journal of Economics* 118, 107–155.

Gu Lifeng, Hackbarth Dirk, 2012, "Governance and Equity Prices: Does Transparency Matter? ," *Review of Finance*, Forthcoming.

Lang, Mark H., and Russel J. Lundholm, 1996, "Corporate disclosure policy and analyst behavior," *Accounting Review* 71, 467–492.

Lang, Mark H., Karl V. Lins, and Mark Maffett, 2010, "Transparency, liquidity, and valuation: international evidence," Working paper, University of Pennsylvania.

Leuz, Christian, and Robert E. Verrecchia, 2000, "The economic consequences of increased disclosure," *Journal of Accounting Research* 38, 91–124.

Leuz, Christian, and Felix Oberholzer, 2006, "Political relationships, global financing, and corporate transparency: evidence from Indonesia," *Journal of Financial Economics* 81, 411–439.

Lim, Terence, 2001, "Rationality and analysts' forecast bias," *Journal of Finance* 56, 369–385.

Martin, Xiumin, and Ron Shalev, 2011, "Target firm-specific information and expected synergies in acquisitions, Working paper, Washington University in St. Louis.

McNichols, Maureen F., 2002, "Discussion of the quality of accruals and earnings: the role of accrual estimation errors," *The Accounting Review* 77 (Supplement), 61–69.

McNichols, Maureen F., and Stephen Stubben, 2011, "The effect of target-firm accounting quality on valuation in acquisitions," Working paper, Stanford University.

Zhang, X. Frank, 2006, "Information uncertainty and stock returns," *Journal of Finance* 61, 105–137.

3.7 Tables and Figures

Table 3.1 : Summary Statistics

This table presents the summary statistics of the independent variables used in the logit estimation model. Q is the ratio of the market value of assets to the book value of assets. PPE is property, plant and equipment scaled by assets. $Cash$ is the ratio of cash and short-term investments to assets. $Size$ is the natural logarithm of firm market capitalization. $Leverage$ is the book debt scaled by assets. ROA is the return on assets. $Industry$ is the dummy variable which equals one if in the previous year there was at least one takeover event in the firm's industry which is defined based on Fama-French 48 industry classifications. $Block$ is also a dummy variable that equals 1 if there is at least one institutional owner whose ownership stake in firm's outstanding shares exceeds 5% and zero otherwise. *Opacity* is measured by accural quality constructed according to McNichols (2002). The details are provided in the data section. Panel A uses all completed takeovers and Panel B uses 100% completed takeovers. The sample period is from 1991 to 2009. t-stat refers to the t-statistic of the difference of the mean between target and non-target group.

	Non-Targets			Targets			
Variable	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.	t-stat
Panel A: Using all completed takeovers							
Q	2.329	1.457	4.383	1.967	1.349	2.375	8.675
PPE	0.533	0.415	1.538	0.509	0.370	0.943	1.421
$Cash$	0.183	0.088	0.221	0.184	0.086	0.222	0.330
$Size$	5.210	5.090	2.226	5.057	4.980	1.920	4.685
$Leverage$	0.175	0.105	0.259	0.201	0.124	0.238	6.492
ROA	-0.071	0.027	0.743	-0.061	0.023	0.356	1.649
$Industry$	0.904	1.000	0.295	0.945	1.000	0.228	10.543
$Block$	0.600	1.000	0.490	0.683	1.000	0.465	10.537
$Opacity$	0.045	0.025	0.245	0.032	0.025	0.025	11.662
Obs. of Non-Targets: 66,265							
Obs. of Targets: 2,420							
Panel B: Using 100% completed takeovers							
Q	2.326	1.455	4.374	2.039	1.417	2.412	5.996
PPE	0.533	0.415	1.542	0.488	0.362	0.459	4.463
$Cash$	0.183	0.087	0.221	0.196	0.096	0.229	3.046
$Size$	5.205	5.086	2.226	5.208	5.118	1.865	0.101
$Leverage$	0.175	0.105	0.259	0.189	0.112	0.225	3.242
ROA	-0.071	0.027	0.741	-0.053	0.027	0.361	2.496
$Industry$	0.853	1.000	0.354	0.928	1.000	0.259	14.803
$Block$	0.600	1.000	0.490	0.730	1.000	0.444	15.291
$Opacity$	0.045	0.025	0.245	0.032	0.025	0.025	11.597
Obs. of Non-Targets: 66,456							
Obs. of Targets: 2,229							

Table 3.2: Logit Estimation of Takeover Likelihood

This table presents the results of the logit regression. The dependent variable is a dummy variable that equals one if the firm is a takeover target in that year. The vector of independent variables includes Q, PPE, $Cash$, $Size$, $Leverage$, ROA, $Industry$, $Block$ and $Opacity$. Q is the ratio of the market value of assets to the book value of assets. PPE is property, plant and equipment scaled by assets. $Cash$ is the ratio of cash and short-term investments to assets. $Size$ is the natural logarithm of firm market capitalization. $Leverage$ is the book debt scaled by assets. ROA is the return on assets. $Industry$ is the dummy variable which equals one if in the previous year there was at least one takeover event in the firm's industry which is defined based on Fama-French 48 industry classifications. $Block$ is also a dummy variable that equals 1 if there is at least one institutional owner whose ownership stake in firm's outstanding shares exceeds 5% and zero otherwise. $Opacity$ is measured by accural quality constructed according to McNichols (2002). All the explanatory variables are measured one year prior to the event year. Results for both all completed takeovers and 100% completed takeovers are reported. The sample period is from 1991 to 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance level 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Model 1		Model 2	
Indep. variable	All completed deals	100% completed deals	All completed deals	100% completed deals
Q	-0.087***	-0.090***	-0.057***	-0.061***
	(5.07)	(4.95)	(3.38)	(3.39)
PPE	0.049	0.034	0.032	0.020
	(1.60)	(1.14)	(0.85)	(0.53)
$Cash$	-0.116	-0.094	-0.082	-0.057
	(0.90)	(0.69)	(0.61)	(0.41)
$Size$	-0.066***	-0.054***	-0.105***	-0.093***
	(6.33)	(4.99)	(9.65)	(8.17)
$Leverage$	0.321***	0.312***	0.307***	0.312***
	(4.16)	(3.34)	(3.82)	(3.34)
ROA	0.206***	0.229***	0.189***	0.298***
	(3.67)	(3.86)	(3.31)	(3.11)
$Industry$	0.618***	0.583***	0.671***	0.634***
	(5.57)	(5.49)	(6.05)	(5.97)
$Block$	0.567***	0.599***	0.549***	0.581***
	(11.34)	(11.47)	(10.98)	(11.09)
$Opacity$			-7.468***	-7.273***
			(10.21)	(9.58)
Observations	68,685	68,685	68,685	68,685
Targets	2,420	2,229	2,420	2,229
Pseudo-R^2	3.41%	3.65%	4.20%	4.38%

Table 3.3: Logit Estimation of Takeover Likelihood: Alternative Transparency Measures
This table presents the results of the logit regression using alternative transparency measures which are constructed from the analyst earnings forecasts. The dependent variable is a dummy variable that equals one if the firm is a takeover target in that year. The vector of independent variables includes Q, PPE, $Cash$, $Size$, $Leverage$, ROA, $Industry$, $Block$ and $Opacity$. Q is the ratio of the market value of assets to the book value of assets. PPE is property, plant and equipment scaled by assets. $Cash$ is the ratio of cash and short-term investments to assets. $Size$ is the natural logarithm of firm market capitalization. $Leverage$ is the book debt scaled by assets. ROA is the return on assets. $Industry$ is the dummy variable which equals one if in the previous year there was at least one takeover event in the firm's industry which is defined based on Fama-French 48 industry classifications. $Block$ is also a dummy variable that equals 1 if there is at least one institutional owner whose ownership stake in firm's outstanding shares exceeds 5% and zero otherwise. $Opacity$ is measured by analyst varaibles: forecast error and forecast dispersion. Forecast error is defined as the absolute value of the difference between the actual annual earnings per share (EPS) and the mean of analyst forecasts. Forecast dispersion is defined as the forecast standard deviation across all analysts following the same firm in the same year. All the explanatory variables are measured one year prior to the event year. Results for both all completed takeovers and 100% completed takeovers are reported. The sample period is from 1991 to 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance level 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

	Forecast Error		Forecast Dispersion	
Indep. variable	All completed deals	100% completed deals	All completed deals	100% completed deals
Q	–0.039***	-0.044***	-0.039***	-0.044***
	(2.85)	(3.02)	(2.87)	(3.05)
PPE	–0.098	–0.082	–0.098	–0.081
	(1.36)	(1.10)	(1.35)	(1.09)
$Cash$	0.031	0.037	0.048	0.054
	(0.25)	(0.29)	(0.39)	(0.42)
$Size$	–0.171***	–0.168***	–0.167***	–0.164***
	(12.90)	(12.28)	(12.69)	(12.08)
$Leverage$	0.524***	0.477***	0.522***	0.473***
	(5.38)	(4.67)	(5.34)	(4.62)
ROA	–0.111*	–0.075	–0.092	–0.053
	(1.72)	(1.03)	(1.40)	(0.70)
$Industry$	0.471***	0.497***	0.474***	0.502***
	(4.16)	(4.62)	(4.20)	(4.66)
$Block$	0.188***	0.194***	0.191***	0.196***
	(3.90)	(3.89)	(3.95)	(3.94)
$Opacity$	–0.378***	–0.388***	–0.320***	-0.327**
	(3.27)	(3.18)	(2.62)	(2.53)
Observations	60,402	60,402	60,402	60,402
Targets	2,772	2,607	2,772	2,607
Pseudo-R^2	4.17%	4.26%	4.11%	4.20%

Table 3.4: Predicted Takeover Probability and Real Takeover Activity

This table reports average predicted takeover likelihood estimated from the logit model 2 and the real takeover activity. Each year the sample used in the logit estimation in Table 3.2 are sorted into quintiles or deciles based on the predicted takeover likelihood. Predicted takeover likelihood is computed using estimation coefficients in Table 3.2 and the logit equation (2). Then I calculate the average predicted takeover probability for each quintile or decile as the mean value of the takeover probability of firms within the same quintile or decile. The realized takeover activity is computed as the takeover event rate within each quintile or decile each year. The average value of predicted takeover probability and the real takeover rate over the sample period of 1991 to 2009 are reported in the table. Panel A and Panel B reports the results for the quintile portfolio and decile portfolio, respectively. The results using model 1 and model 2 are shown in the right and left part of the table respectively. Model 1 refer to the old logit model and Model 2 refers to the new logit model with transparency as the additional variable.

	Model 2		Model 1	
	Predicted Takeover Likelihood	Real Takeover Rate	Predicted Takeover Likelihood	Real Takeover Rate
	Panel A: Quintile Portfolio			
Low	0.0152	0.0117	0.0180	0.0165
2	0.0257	0.0295	0.0270	0.0212
3	0.0336	0.0360	0.0322	0.0384
4	0.0409	0.0394	0.0403	0.0437
High	0.0527	0.0515	0.0487	0.0476
	Panel B: Decile Portfolio			
Low	0.0112	0.0077	0.0149	0.0114
2	0.0192	0.0156	0.0211	0.0227
3	0.0236	0.0276	0.0251	0.0261
4	0.0278	0.0315	0.0289	0.0283
5	0.0318	0.0346	0.0325	0.0309
6	0.0355	0.0373	0.0358	0.0340
7	0.0390	0.0369	0.0388	0.0357
8	0.0429	0.0419	0.0418	0.0458
9	0.0478	0.0501	0.0453	0.0543
High	0.0575	0.0529	0.0521	0.0469

Table 3.5: Takeover likelihood and equity return relation

This table reports the monthly returns and abnormal returns to the quintile portfolios formed based on firms' takeover likelihood, which is constructed using the logit estimation coefficients in Table 3.2. Results for all completed deals sample and 100% completed deals sample are presented in Panel A and Panel B, respectively. In each Panel, the equal-weighted mean return to the quintile portfolio, both equal-weighted and value-weighted return spread and abnormal returns to the quintile or decile hedge portfolio which buys firms with the highest takeover probability and sells firms with the lowest takeover probability, the return and abnormal return to the long-short portfolio when the takeover probability is constructed using the logit estimation of the old model are also shown at the bottom of each panel for compasion. In order to correct the in-sample "look-ahead" bias, I estimate the logit model 2 using 10-year rolling window and report the portfolio return results in the right part of the table. The sample period is from 1991 to 2009. t-statistics are reported after the estimation coefficient.

	Whole sample logit estimation				Rolling windows logit estimation			
Takeover prob	Mean	t	Alpha	t	Mean	t	Alpha	t
Panel A: All completed deals sample								
Low	1.06%	2.93	0.09%	0.84	0.72%	1.15	0.37%	1.96
2	1.32%	3.39	0.39%	2.46	1.10%	1.77	0.68%	2.88
3	1.58%	4.31	0.54%	4.55	1.16%	2.05	0.58%	3.82
4	1.41%	3.61	0.28%	2.85	1.17%	1.90	0.35%	1.98
High	2.02%	4.74	0.97%	4.83	1.95%	2.90	1.17%	3.77
H-L(EW)	0.96%	4.81	0.88%	5.14	1.23%	4.24	0.80%	3.22
H-L(VW)	0.84%	3.03	0.42%	2.55	1.20%	3.11	0.35%	1.37
H-L(EW, decile)	1.47%	5.25	1.40%	5.79	1.85%	4.43	1.26%	3.52
H-L(VW, decile)	1.18%	3.47	0.68%	3.33	1.58%	3.22	0.62%	1.66
H-L(EW)	0.89%	4.73	0.79%	5.01	1.12%	3.81	0.69%	2.82
H-L(VW)	0.67%	2.62	0.21%	1.89	0.95%	2.60	0.35%	1.24
H-L(EW, decile)	1.40%	5.31	1.27%	5.67	1.60%	3.66	1.05%	2.91
H-L(VW, decile)	0.93%	3.84	0.55%	2.76	1.44%	2.83	0.53%	1.87
Panel B: 100% completed deals sample								
Low	1.08%	2.93	0.11%	0.97	0.76%	1.14	0.42%	1.93
2	1.48%	3.63	0.55%	3.10	1.23%	1.99	0.73%	2.99
3	1.56%	4.28	0.52%	4.10	1.13%	1.98	0.59%	4.04
4	1.32%	3.48	0.20%	4.04	1.09%	1.82	0.29%	1.58
High	1.94%	4.59	0.89%	4.62	1.86%	2.80	1.09%	3.72
H-L(EW)	0.85%	4.61	0.78%	4.84	1.10%	3.82	0.68%	2.80
H-L(VW)	0.71%	2.68	0.35%	2.05	0.94%	2.59	0.39%	1.31
H-L(EW, decile)	1.36%	5.20	1.26%	5.52	1.58%	3.75	1.06%	2.99
H-L(VW, decile)	1.06%	3.38	0.63%	3.03	1.18%	2.43	0.41%	1.52
H-L(EW)	0.82%	4.67	0.73%	4.84	0.95%	3.11	0.53%	2.16
H-L(VW)	0.67%	2.78	0.32%	2.15	0.76%	2.18	0.19%	0.69
H-L(EW, decile)	1.27%	5.24	1.14%	5.52	1.43%	3.35	0.89%	2.58
H-L(VW, decile)	0.72%	2.54	0.34%	1.68	1.09%	2.29	0.34%	1.09

Table 3.6: Summary statistics of factors

This table lists summary statistics of the takeover factor, Fama-French three factors, and Carhart momentum factor constructed according to Carhart (1997). The five factors are denoted by TOP(takeover factor), MKT(market factor), SML(size factor), HML(value factor), and UMD(momentum factor), respectively. The takeover factor is constructed as the monthly equal-weighted portfolio return to the hedge portfolio that is long in firms in the top takeover likelihood quintile and short in firms from the bottom takeover likelihood quintile. Panel A lists some basic statistics of the five factors. $SKEW$ and $KURT$ refer to skewness and kurtosis, respectively. Panel B lists the correlation matrix of these factors. The time period is from 1991 to 2009.

Panel A: Basic descriptive statistics					
	Mean	T–stat	STD	SKEW	KURT
MKT	0.55	1.86	4.43	-0.85	1.80
SMB	0.27	1.15	3.57	0.81	7.85
HML	0.37	1.64	3.42	0.00	2.56
UMD	0.39	0.87	6.81	-1.37	6.78
TOP	0.96	4.81	3.01	0.72	2.25
Panel B: Correlation matrix of factors					
	MKT	SMB	HML	UMD	TOP
MKT	1.00				
SMB	0.22	1.00			
HML	-0.28	-0.37	1.00		
UMD	-0.34	-0.01	-0.04	1.00	
TOP	-0.09	0.32	0.21	-0.26	1.00

Table 3.7: Pricing the 25 Size and B/M portfolios: Equal-Weighted

This table reports the intercept α from the time-series regressions of monthly size and B/M portfolio returns on risk factors using different asset-pricing models. Panel A presents the mean excess return of each portfolio. Panel B presents the α_s for the market model. Panel C reports intercept α_s for the model with CAPM and the new takeover factor. Panel D and Panel E report the α_s for Fama-French four-factor model and the augmented five-factor model. The right part of each panel reports the t-statistics of the estimates. The data of the 25 size and B/M portfolio returns, Fama-French risk factors, risk-free rates are from Kenneth-French's website. The momentum factor are constructed according to the procedure in Carhart (1997). The sample period is from 1991 to 2009.

	Book-to-market ratio									
Size	Low	2	3	4	High	Low	2	3	4	High
	Panel $A:\ R_t - R_f$									
	α					t-stat				
Small	0.52	1.14	1.25	1.25	1.69	0.83	2.35	3.22	3.60	4.11
2	0.46	0.79	1.10	0.93	1.06	0.87	1.97	3.13	2.59	2.42
3	0.55	0.81	0.98	0.95	1.32	1.10	2.17	2.91	2.76	3.48
4	0.72	0.79	0.82	0.90	0.75	1.67	2.30	2.30	2.62	1.91
Big	0.55	0.72	0.68	0.73	0.85	1.56	2.31	2.10	2.28	2.42
	Panel $B:\ R_t - R_f = \alpha + \gamma RMRF_t + \epsilon_t$									
	α					t-stat				
Small	-0.26	0.48	0.72	0.79	1.17	0.55	1.44	2.72	3.23	3.86
2	-0.33	0.16	0.56	0.39	0.43	1.06	0.75	2.79	1.81	1.53
3	-0.23	0.20	0.45	0.43	0.77	0.85	1.11	2.52	2.11	3.20
4	0.00	0.23	0.26	0.37	0.18	0.02	1.39	1.36	1.91	0.72
Big	-0.07	0.20	0.17	0.26	0.39	0.63	1.44	0.98	1.32	1.55
	Panel $C:\ R_t - R_f = \alpha + \beta_1 RMRF_t + \gamma TOP_t + \epsilon_t$									
	α					t-stat				
Small	-1.44	-0.58	-0.21	-0.11	0.01	3.39	2.10	1.07	0.62	0.04
2	-0.88	-0.31	0.11	-0.08	-0.26	2.82	1.50	0.59	0.39	0.99
3	-0.58	-0.10	0.19	0.11	0.42	2.08	0.55	1.08	0.54	1.72
4	-0.13	0.03	-0.03	0.16	-0.18	0.63	0.16	0.15	0.82	0.74
Big	-0.08	0.10	0.10	0.08	0.16	0.68	0.67	0.58	0.41	0.63
	Panel $D: R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \epsilon_t$									
	α					t-stat				
Small	-0.05	0.48	0.56	0.55	0.89	0.18	2.86	4.30	4.19	5.06
2	-0.18	0.00	0.23	-0.04	-0.15	1.32	0.04	2.06	0.38	1.22
3	-0.02	0.03	0.14	0.03	0.34	-0.10	0.20	1.09	0.22	2.32
4	0.22	0.03	0.00	0.04	-0.22	1.66	0.18	-0.02	0.26	1.37
Big	0.14	0.10	-0.02	-0.03	0.05	1.69	0.87	-0.16	-0.25	0.33
	Panel E: $R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \gamma TOP_t + \epsilon_t$									
	α					t-stat				
Small	-0.39	0.10	0.19	0.13	0.31	1.54	0.62	1.57	1.17	2.03
2	-0.11	0.06	0.27	0.00	-0.13	0.76	0.50	2.31	0.03	1.03
3	0.06	0.12	0.25	0.07	0.56	0.39	0.80	1.88	0.46	3.65
4	0.30	0.03	-0.03	0.10	-0.20	2.16	0.22	0.18	0.71	1.20
Big	0.13	0.03	-0.01	-0.05	0.10	1.47	0.25	0.05	0.42	0.58

Table 3.8: Pricing the 25 Size and B/M portfolios: value-weighted

This table reports the intercept α from the time-series regressions of monthly size and B/M portfolio returns on risk factors using different asset-pricing models. Panel A presents the mean excess return of each portfolio. Panel B presents the α_s for the market model. Panel C reports intercept α_s for the model with CAPM and the new takeover factor. Panel D and Panel E report the α_s for Fama-French four-factor model and the augmented five-factor model. The right part of each panel reports the t-statistics of the estimates. The data of the 25 size and B/M portfolio returns, Fama-French risk factors, risk-free rates are from Kenneth-French's website. The sample period is from 1991 to 2009.

	Book-to-market ratio									
Size	Low	2	3	4	High	Low	2	3	4	High
	Panel A : $R_t - R_f$									
	α					t-stat				
Small	0.19	0.99	1.00	1.17	1.32	0.33	2.07	2.65	3.31	3.38
2	0.48	0.72	1.01	0.92	1.01	0.98	1.92	3.05	2.70	2.52
3	0.49	0.76	0.92	0.89	1.25	1.06	2.19	2.98	2.80	3.61
4	0.70	0.76	0.75	0.88	0.74	1.72	2.37	2.26	2.82	2.10
Big	0.53	0.66	0.51	0.53	0.65	1.73	2.35	1.75	1.78	1.86
	Panel B : $R_t - R_f = \alpha + \gamma RMRF_t + \epsilon_t$									
	α					t-stat				
Small	-0.60	0.35	0.47	0.68	0.77	1.53	1.04	1.88	2.84	2.98
2	-0.26	0.14	0.51	0.42	0.44	0.91	0.68	2.62	1.99	1.69
3	-0.23	0.19	0.43	0.42	0.75	0.91	1.17	2.62	2.16	3.40
4	0.03	0.23	0.22	0.39	0.23	0.15	1.58	1.28	2.30	1.02
Big	0.00	0.20	0.05	0.11	0.18	0.02	1.52	0.31	0.55	0.74
	Panel C : $R_t - R_f = \alpha + \beta_1 RMRF_t + \gamma TOP_t + \epsilon_t$									
	α					t-stat				
Small	-1.26	-0.31	-0.16	0.06	-0.05	3.25	0.97	0.70	0.30	0.26
2	-0.60	-0.21	0.19	0.05	-0.07	2.02	0.98	0.97	0.23	0.27
3	-0.39	0.00	0.28	0.20	0.53	1.48	0.02	1.60	1.01	2.34
4	0.00	0.08	0.04	0.28	0.01	0.01	0.57	0.22	1.57	0.06
Big	0.12	0.24	-0.01	0.03	0.00	0.10	1.76	0.04	0.15	0.02
	Panel D :$R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \epsilon_t$									
	α					t-stat				
Small	-0.57	0.13	0.17	0.29	0.30	3.35	0.98	1.62	2.66	2.85
2	-0.26	-0.10	0.13	-0.04	-0.15	2.28	0.95	1.34	0.41	1.43
3	-0.14	-0.02	0.11	0.01	0.31	1.19	0.12	0.90	0.07	2.14
4	0.13	0.02	-0.06	0.05	-0.18	1.23	0.12	0.46	0.43	1.26
Big	0.22	0.14	-0.09	-0.16	-0.10	2.86	1.32	0.82	1.47	0.59
	Panel E: $R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \gamma TOP_t + \epsilon_t$									
	α					t-stat				
Small	-0.64	-0.01	-0.05	0.11	0.02	3.52	0.08	0.47	1.02	0.16
2	-0.20	-0.05	0.21	0.02	-0.07	1.65	0.43	2.05	0.18	0.62
3	-0.07	0.06	0.21	0.05	0.54	0.57	0.45	1.60	0.37	3.65
4	0.19	0.03	-0.07	0.14	-0.13	1.60	0.26	0.46	1.08	0.83
Big	0.15	0.10	-0.12	-0.16	-0.15	1.82	0.93	0.97	1.41	0.81

Table 3.9: Pricing the 25 Size and B/M portfolios: Equal-Weighted (old factor)

This table reports the intercept α from the time-series regressions of monthly equal-weighted size and B/M portfolio returns on risk factors using different asset-pricing models. Panel A presents the mean excess return of each portfolio. Panel B presents the α_s for the Carhart (1997) four-factor model. Panel C reports intercept α_s for the augmented five-factor model including the new takeover factor TOP as the fifth pricing factor and Panel D reports intercept α_s for the augmented five-factor model including the old takeover factor as the fifth pricing factor (The old takeover factor $TOPO$ is constructed using the old logit model in the prior literature.) The right part of each panel reports the t-statistics of the estimates. The data of the 25 size and B/M portfolio returns, Fama-French risk factors, risk-free rates are from Kenneth-French's website. The momentum factor is constructed according to the procedure in Carhart (1997). The sample period is from 1991 to 2009.

	Book-to-market ratio									
Size	Low	2	3	4	High	Low	2	3	4	High
	Panel $A: R_t - R_f$									
	α					t-stat				
Small	0.52	1.14	1.25	1.25	1.69	0.83	2.35	3.22	3.60	4.11
2	0.46	0.79	1.10	0.93	1.06	0.87	1.97	3.13	2.59	2.42
3	0.55	0.81	0.98	0.95	1.32	1.10	2.17	2.91	2.76	3.48
4	0.72	0.79	0.82	0.90	0.75	1.67	2.30	2.30	2.62	1.91
Big	0.55	0.72	0.68	0.73	0.85	1.56	2.31	2.10	2.28	2.42
	Panel $B: R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \epsilon_t$									
	α					t-stat				
Small	-0.05	0.48	0.56	0.55	0.89	0.18	2.86	4.30	4.19	5.06
2	-0.18	0.00	0.23	-0.04	-0.15	1.32	0.04	2.06	0.38	1.22
3	-0.02	0.03	0.14	0.03	0.34	-0.10	0.20	1.09	0.22	2.32
4	0.22	0.03	0.00	0.04	-0.22	1.66	0.18	-0.02	0.26	1.37
Big	0.14	0.10	-0.02	-0.03	0.05	1.69	0.87	-0.16	-0.25	0.33
	Panel C: $R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \gamma TOP_t + \epsilon_t$									
	α					t-stat				
Small	-0.39	0.10	0.19	0.13	0.31	1.54	0.62	1.57	1.17	2.03
2	-0.11	0.06	0.27	0.00	-0.13	0.76	0.50	2.31	0.03	1.03
3	0.06	0.12	0.25	0.07	0.56	0.39	0.80	1.88	0.46	3.65
4	0.30	0.03	-0.03	0.10	-0.20	2.16	0.22	0.18	0.71	1.20
Big	0.13	0.03	-0.01	-0.05	0.10	1.47	0.25	0.05	0.42	0.58
	Panel D: $R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \gamma TOPO_t + \epsilon_t$									
	α					t-stat				
Small	-0.33	0.15	0.23	0.17	0.39	1.64	0.66	1.62	1.24	2.67
2	-0.06	0.08	0.28	0.02	-0.14	0.76	0.50	2.35	0.05	1.02
3	0.07	0.12	0.25	0.06	0.58	0.41	0.89	1.87	0.43	3.69
4	0.35	0.05	-0.04	0.13	-0.24	2.26	0.28	0.38	0.71	1.23
Big	0.17	0.04	-0.04	-0.03	0.15	1.53	0.27	0.04	0.40	0.58

Table 3.10: Pricing the 25 Size and B/M portfolios: value-weighted (old factor)

This table reports the intercept α from the time-series regressions of monthly value-weighted size and B/M portfolio returns on risk factors using different asset-pricing models. Panel A presents the mean excess return of each portfolio. Panel B presents the α_s for the Carhart (1997) four-factor model. Panel C reports intercept α_s for the augmented five-factor model including the new takeover factor TOP as the fifth pricing factor and Panel D reports intercept α_s for the augmented five-factor model including the old takeover factor $TOPO$ as the fifth pricing factor (The old takeover factor is constructed using the old logit model in the literature.) The right part of each panel reports the t-statistics of the estimates. The data of the 25 size and B/M portfolio returns, Fama-French risk factors, risk-free rates are from Kenneth-French's website. The momentum factor is constructed according to the procedure in Carhart (1997). The sample period is from 1991 to 2009.

	Book-to-market ratio									
Size	Low	2	3	4	High	Low	2	3	4	High
	Panel A : $R_t - R_f$									
	α					t-stat				
Small	0.19	0.99	1.00	1.17	1.32	0.33	2.07	2.65	3.31	3.38
2	0.48	0.72	1.01	0.92	1.01	0.98	1.92	3.05	2.70	2.52
3	0.49	0.76	0.92	0.89	1.25	1.06	2.19	2.98	2.80	3.61
4	0.70	0.76	0.75	0.88	0.74	1.72	2.37	2.26	2.82	2.10
Big	0.53	0.66	0.51	0.53	0.65	1.73	2.35	1.75	1.78	1.86
	Panel D :$R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \epsilon_t$									
	α					t-stat				
Small	-0.57	0.13	0.17	0.29	0.30	3.35	0.98	1.62	2.66	2.85
2	-0.26	-0.10	0.13	-0.04	-0.15	2.28	0.95	1.34	0.41	1.43
3	-0.14	-0.02	0.11	0.01	0.31	1.19	0.12	0.90	0.07	2.14
4	0.13	0.02	-0.06	0.05	-0.18	1.23	0.12	0.46	0.43	1.26
Big	0.22	0.14	-0.09	-0.16	-0.10	2.86	1.32	0.82	1.47	0.59
	Panel E: $R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \gamma TOP_t + \epsilon_t$									
	α					t-stat				
Small	-0.64	-0.01	-0.05	0.11	0.02	3.52	0.08	0.47	1.02	0.16
2	-0.20	-0.05	0.21	0.02	-0.07	1.65	0.43	2.05	0.18	0.62
3	-0.07	0.06	0.21	0.07	0.54	0.57	0.45	1.60	0.37	3.65
4	0.19	0.03	-0.07	0.14	-0.13	1.60	0.26	0.46	1.08	0.83
Big	0.15	0.10	-0.12	-0.16	-0.15	1.82	0.93	0.97	1.41	0.81
	Panel D: $R_t - R_f = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_2 UMD_t + \gamma TOPO_t + \epsilon_t$									
	α					t-stat				
Small	-0.60	0.01	0.07	0.16	0.08	3.59	0.18	0.51	1.05	1.20
2	-0.20	-0.02	0.28	0.09	-0.01	1.75	0.44	2.16	0.18	0.70
3	-0.02	0.08	0.25	0.05	0.56	0.52	0.49	1.70	0.37	3.85
4	0.23	0.06	-0.07	0.17	-0.14	1.65	0.36	0.66	1.20	0.85
Big	0.15	0.14	-0.10	-0.12	-0.11	1.88	0.93	0.87	1.21	0.71

Table 3.11: Premium associated with the takeover exposure

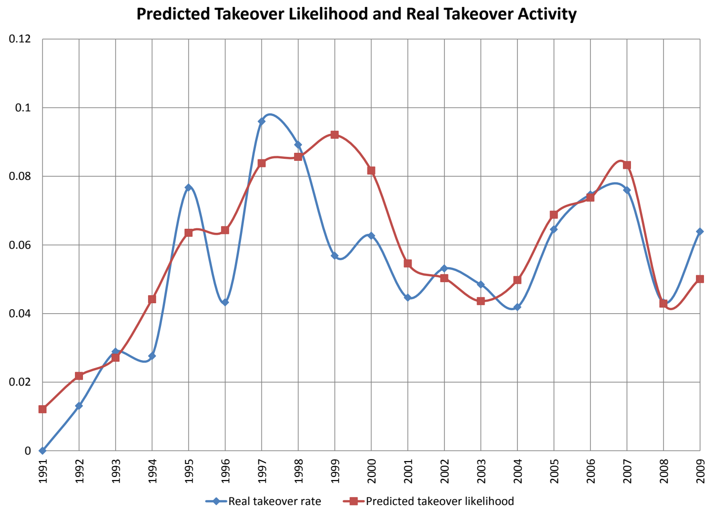
This table reports the coefficients of the regression of mean excess return of each of the 100 size and book-to-market sorted portfolios on the portfolio betas which are computed as the coefficient of the regression of the excess return of each of the 100 portfolios on factors. TOP refers to the takeover factor constructed using the logit estimation results from model 2 and $TOPO$ refers to the takeover factor constructed using the logit estimation results from model 1. The data of the 100 size and Book-to-market portfolio returns, Fama-French risk factors, risk-free rates are from Kenneth-French's website. The momentum factor is constructed according to the procedure in Carhart (1997). The sample period is from 1991 to 2009.

Panel A: Pricing with Carhart (1997) Four-Factor Model						
	FF4	*t*-stat	FF4 + TOPO	*t*-stat	FF4 +TOP	*t*-stat
Intercept	0.20	(7.49)	0.16	(5.99)	0.18	(7.65)
MARKET	-0.12	(4.49)	-0.08	(3.14)	-0.10	(4.78)
SIZE	0.03	(3.80)	0.04	(4.35)	0.04	(4.16)
HML	0.04	(5.37)	0.04	(6.18)	0.04	(5.55)
UMD	0.04	(1.19)	0.04	(1.40)	0.04	(1.26)
TOPO			0.06	(4.23)		
TOP					0.07	(3.93)
R^2	0.21		0.41		0.50	

Panel B: Pricing with Capital Asset Pricing Model (CAPM)						
	CAPM	*t*-stat	CAPM + TOPO	*t*-stat	CAPM + TOP	*t*-stat
Intercept	0.21	(6.77)	0.18	(5.42)	0.22	(8.47)
MARKET	-0.10	(3.34)	-0.09	(2.27)	-0.11	(4.59)
TOPO			0.05	(2.55)		
TOP					0.08	(4.05)
R^2	0.10		0.15		0.21	

Figure 3.1: Predicted takeover likelihood and real takeover activity

This figure plots the time series of the average predicted takeover likelihood and the real takeover rates for the top decile group over the period of 1991 to 2009. Predicted takeover likelihood is computed using estimation coefficients in Table 3.2 and the logit equation (2). Each year firms are sorted into deciles based on the predicted takeover probability. Then we calculate the average predicted takeover probability for each decile as the mean value of the takeover probability of firms within the same decile. The realized takeover activity is computed as the takeover event rate within each decile. The time series of the average predicted takeover likelihood and the real takeover rates for decile 10 (with highest predicted takeover likelihood)are plotted in the figure.



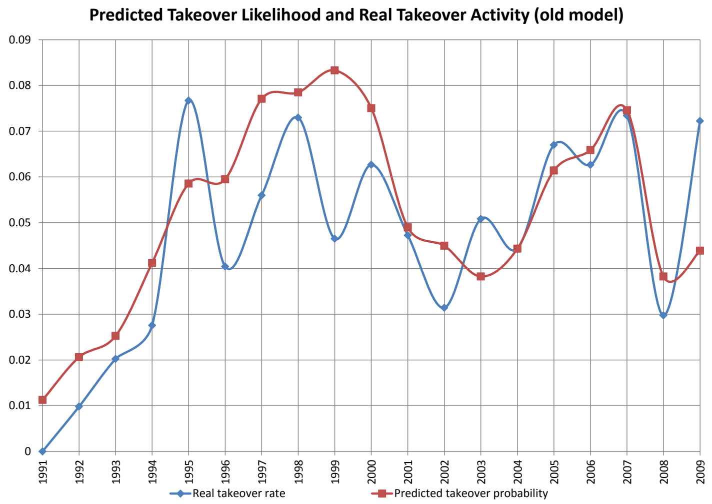
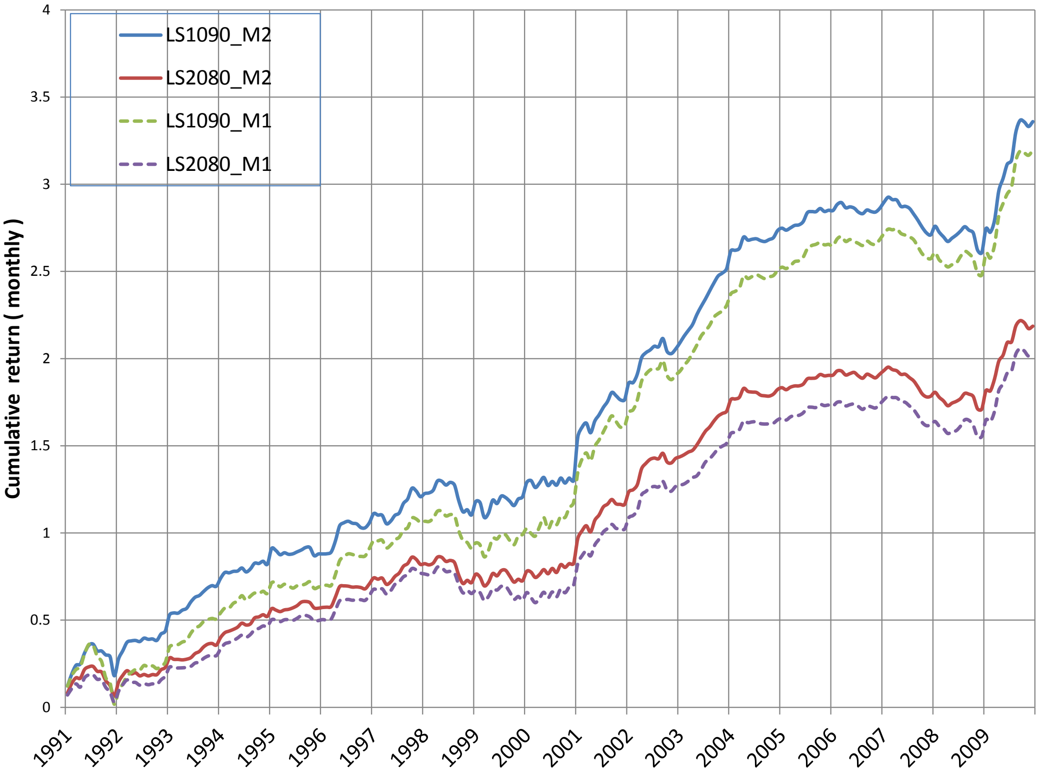


Figure 3.2: Cumulative return of the long-short portfolio

This figure plots the time series of the cumulative monthly return for the long-short portfolio formed based on the predicted takeover probability over the period of 1991 to 2009. LS1090 refers to the portfolio that buys stocks in the top 10% takeover likelihood group and sells stocks in the bottom 10% takeover likelihood group. Similarly, LS2080 refers to the long-short protfolio sorted in quintiles. M1 refers to the portfolio sorted based on the takeover likelihood using the resutls from model 1 and M2 refers to the portfolio sorted based on the takeover likelihood using the estiamtion results from model 2.



Appendix A

A.1 Solution of the valuation function $V(c, n)$

For $n < N$, let $V^c(c, n)$ solve equation (6) with $u = 1$ and $V^m(c, n)$ solve equation (6) with $u = 0$. Let these functions satisfy the single crossing property in c. Then for every n, there exists a threshold c_n^* such that

$$V(c, n) = \begin{cases} V^c(c, n) = \sum_{i=n}^{N-1} F_n^i (logc)^{i-n} + B_n c + G_n & c \geq c_n^* \\ V^m(c, n) = A_n c^\beta & c < c_n^* \end{cases} \tag{A.1}$$

with

$$V(0, n) = 0 \tag{A.2}$$

$$\lim_{c \to \infty} \frac{V(c, n)}{c} < \infty \tag{A.3}$$

When $c = c_n^*$, the functions satisfy the following boundary conditions:

$$V^c(c_n^*, n) = V^m(c_n^*, n) \tag{A.4}$$

$$\frac{\partial}{\partial c} V^c(c_n^*, n) = \frac{\partial}{\partial c} V^m(c_n^*, n) \tag{A.5}$$

$$\frac{\partial^2}{\partial c^2} V^c(c_n^*, n) = \frac{\partial^2}{\partial c^2} V^m(c_n^*, n) \tag{A.6}$$

Since the cash flow process follows a geometric Brownian process once the cash flow reaches zero it will stay there forever, the firm value is zero when $c = 0$. That is, $V(0, n) = 0$. When the cash flow approaches infinity, in order to prevent bubbles, firm value increases in proportion to future cash flows. That is, $\lim_{c \to \infty} \frac{V(c,n)}{c} < \infty$.

When $u = 0$, the Hamilton-Bellman-Jacobi equation of the investment problem is reduced to the following ordinary differential equation:

$$\hat{r} V(c, n) = \frac{1}{2} \sigma^2 c^2 \frac{\partial^2}{\partial c^2} V(c, n) + \hat{\mu} c \frac{\partial}{\partial c} V(c, n) \tag{A.7}$$

The solution of this HJB equation has a standard solution given as follows:

$$V^m(c, n) = A_n c^\beta \tag{A.8}$$

Substitute the value function $V^m(c,n)$ into the HJB equation will give the value of β:

$$\beta = \frac{(\sigma^2 - 2\hat{\mu}) + \sqrt{8(r+\phi)\sigma^2 + (\sigma^2 - 2\hat{\mu})^2}}{2\sigma^2} \tag{A.9}$$

Substitute the value function $V^c(c,n)$ into the HJB equation when $u = 0$ and applying the three boundary conditions will give the formula of A_n, B_n, G_n, F_n^i and, c_n^*. They are as follows:

$$F_n^i = \begin{cases} 0 & i = N \\ \frac{2\pi}{(i-n)\psi} F_{n+1}^i + \frac{\sigma^2(i-n+1)}{\psi} F_n^{i+1} & n < i < N \\ \frac{2\sigma^2 (c_n^*)^{-\gamma}}{\psi+\theta}((1-\beta)B_n c_n^* - \beta G_n) - \sum_{i=n+1}^{N-1} F_n^i((logc_n^*)^{i-n} - \frac{2\sigma^2(i-n)}{\psi+\theta}(logc_n^*)^{i-n-1}) & i = n \end{cases} \tag{A.10}$$

$$A_n = (B_n c_n^* + G_n + \sum_{i=n}^{N-1} F_n^i (c_n^*)^{\gamma} (logc_n^*)^(i-n))(c_n^*)^{-\beta} \tag{A.11}$$

$$B_n = \frac{B_{n+1}\pi}{r+\phi+\pi-\hat{\mu}} \quad with\ B(N) = \frac{1}{r+\phi-\hat{\mu}} \tag{A.12}$$

$$G_n = \frac{G_{n+1}\pi - a(n)}{r+\phi+\pi} \quad with\ G(N) = 0 \tag{A.13}$$

and c_n^* can be derived by solving the following equation:

$$(\beta-1)(1-\gamma)B_n c_n^* - \beta\gamma G_n = \frac{(c_n^*)^{\gamma}}{2\sigma^2} \sum_{i=n+1}^{N-1} ((i-n)(\psi-\theta)F_n^i - 4\pi F_{n+1}^i)(logc_n^*)^{i-n-1} \tag{A.14}$$

where

$$\psi = \sqrt{8(r+\phi+\pi)\sigma^2 + (\sigma^2 - 2\hat{\mu})^2} \tag{A.15}$$

$$\theta = \sqrt{8(r+\phi)\sigma^2 + (\sigma^2 - 2\hat{\mu})^2} \tag{A.16}$$

$$\gamma = \frac{(\sigma^2 - 2\hat{\mu}) - \psi}{2\sigma^2} \tag{A.17}$$

The risk premium can be computed as

$$\frac{(\partial V(c,n)/\partial c)c}{V(c,n)}\lambda \tag{A.18}$$

Appendix B

B.1 Additional Evidence

Throughout the empirical analysis, portfolios are formed based on tercile breakpoints of the competition measure and R&D intensity measure. In order to make sure that the main results are not sensitive to the particular way of sorting. I perform the tests with various kinds of sorting methods (For example, 3×5 (i.e., tercile on Herfindahl Index, quintile on R&D intensity measure) or 5×3 (i.e., quintile on Herfindahl Index, tercile on R&D intensity measure)). Here in the appendix, I sort firms according to the quintile breakpoints of these two measures and the results are presented in Table B.1 and Table B.2. As is shown in the table, this refined sorting approach produces similar and sometimes stronger results.

B.2 Tables and Figures

Table B.1: Double Sorting on Competition and R&D Intensity: 5×5 **Sort**

This table reports the monthly abnormal returns (in %) of quintile portfolios sorted on product market competition and R&D intensity. In June of year t, NYSE, Amex, and NASDAQ stocks are sorted into quintiles based on the Industry Herfindahl index in year $t-1$. Meanwhile, independently, firms with non-missing R&D are sorted into quintile portfolios based on R&D intensity measure in year $t-1$. Monthly equal-weighted and value-weighted returns on the resulting portfolios are calculated from July of year t to June of year $t+1$. Monthly portfolio abnormal returns are computed by running time-series regression of portfolio excess returns on risk factors. The measure for product market competition is Herfindahl index computed based on the market share of each firm in the industry. The measure for R&D intensity is R&D expenditure scaled by net sales (RDS) or R&D capital scaled by assets ($RDCA$). The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Equal-Weighted Portfolio Return: RDS									
		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
FF 3-factor	α	0.06	–0.09	–1.51**	–1.57**	–0.26**	0.23**	0.39**	0.65***
	t-stat	(0.20)	(0.31)	(2.13)	(2.52)	(2.04)	(2.00)	(2.11)	(3.38)
Carhart 4-factor	α	0.29	0.07	–1.13	–1.42**	–0.06	0.42***	0.61***	0.67***
	t-stat	(0.92)	(0.25)	(1.60)	(2.28)	(0.52)	(4.23)	(3.47)	(3.43)
Panel B: Value-Weighted Portfolio Return : RDS									
		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
FF 3-factor	α	–0.37	–0.06	–1.59**	–1.21**	0.01	0.06	0.41**	0.40**
	t-stat	(1.24)	(0.22)	(2.45)	(1.99)	(0.08)	(0.66)	(2.49)	(1.99)
Carhart 4-factor	α	–0.14	–0.01	–1.26*	–1.29*	0.07	0.07*	0.43***	0.36*
	t-stat	(0.48)	(0.03)	(1.95)	(1.78)	(0.49)	(0.77)	(2.61)	(1.90)

Panel C: Equal-Weighted Portfolio Return: RDCA									
		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
FF 3-factor	α	–0.23	0.05	–1.02	–0.79	0.03	0.22*	0.72***	0.68***
	t-stat	(0.67)	(0.16)	(0.63)	(0.78)	(0.28)	(1.84)	(3.67)	(3.33)
Carhart 4-factor	α	0.00	0.16	–0.45	–0.45	0.14	0.38***	0.90***	0.76***
	t-stat	(0.01)	(0.47)	(0.28)	(0.64)	(1.19)	(3.55)	(4.79)	(3.72)
Panel D: Value-Weighted Portfolio Return : RDCA									
		Low Competition				High Competition			
		$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L	$R\&D_L$	$R\&D_M$	$R\&D_H$	H-L
FF 3-factor	α	–0.61*	–0.26	–1.01	–0.52	0.14	0.16	0.67***	0.54**
	t-stat	(1.77)	(0.80)	(0.60)	(0.29)	(1.07)	(1.54)	(3.97)	(2.41)
Carhart 4-factor	α	–0.42	–0.27	–0.54	–0.31	0.13	0.25**	0.63***	0.49**
	t-stat	(1.02)	(0.85)	(0.00)	(0.44)	(1.04)	(2.48)	(3.67)	(2.20)

Table B.2: Double Sorting on Competition and R&D Intensity: 5×5 Sort

This table reports the monthly abnormal returns (in %) of quintile portfolios sorted on product market competition and R&D intensity. In June of year t, NYSE, Amex, and NASDAQ stocks are sorted into quintiles based on the Industry Herfindahl index in year $t-1$. Meanwhile, independently, firms with non-missing R&D are sorted into quintile portfolios based on R&D intensity measure in year $t-1$. Monthly equal-weighted and value-weighted returns on the resulting portfolios are calculated from July of year t to June of year $t+1$. Monthly portfolio abnormal returns are computed by running time-series regression of portfolio excess returns on risk factors. The measure for product market competition is Herfindahl index computed based on the market share of each firm in the industry. The measure for R&D intensity is R&D expenditure scaled by net sales (RDS) or R&D capital scaled by assets ($RDCA$). The sample period is from July 1963 to December 2009. t-statistics are reported in parentheses under the estimation coefficient. The significance levels 1%, 5%, and 10% are denoted by ***, **, and *, respectively.

Panel A: Equal-Weighted Portfolio Return: RDS

		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
FF 3-factor	α	0.06	–0.23	–0.26**	–0.32*	–1.51**	–0.30	0.39**	1.90**
	t-stat	(0.20)	(1.16)	(2.04)	(1.72)	(2.13)	(0.85)	(2.11)	(2.08)
Carhart 4-factor	α	0.29	–0.10	–0.06	–0.35	–1.13	0.05	0.61***	1.74***
	t-stat	(0.92)	(0.51)	(0.52)	(0.41)	(1.60)	(0.16)	(3.47)	(3.25)

Panel B: Value-Weighted Portfolio Return : RDS

		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
FF 3-factor	α	–0.37	–0.33	0.01	0.47	–1.59**	0.40*	0.41**	2.00**
	t-stat	(1.24)	(1.41)	(0.08)	(0.89)	(2.45)	(1.96)	(2.49)	(1.98)
Carhart 4-factor	α	–0.14	–0.38	0.07	0.21	–1.26*	0.41**	0.43***	1.69*
	t-stat	(0.48)	(1.60)	(0.49)	(0.71)	(1.95)	(2.17)	(2.61)	(1.93)

Panel C: Equal-Weighted Portfolio Return: RDCA

		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
FF 3-factor	α	–0.23	–0.03	0.03	0.26	–1.02	0.70	0.72***	1.74***
	t-stat	(0.67)	(0.15)	(0.28)	(0.38)	(0.63)	(1.61)	(3.67)	(3.17)
Carhart 4-factor	α	0.00	0.07	0.14	0.14	–0.45	0.81**	0.90***	1.35***
	t-stat	(0.01)	(0.34)	(1.19)	(0.69)	(0.28)	(2.13)	(4.79)	(2.95)

Panel D: Value-Weighted Portfolio Return : RDCA

		Low R&D Intensity				High R&D Intensity			
		HHI_H	HHI_M	HHI_L	L-H	HHI_H	HHI_M	HHI_L	L-H
FF 3-factor	α	–0.61*	0.01	0.14	0.75	–1.01	0.34	0.67***	1.68***
	t-stat	(1.77)	(0.06)	(1.07)	(0.99)	(0.60)	(0.70)	(3.97)	(2.89)
Carhart 4-factor	α	–0.42	–0.09	0.13	0.55	–0.54	0.36	0.63***	1.17***
	t-stat	(1.23)	(0.34)	(1.04)	(0.46)	(0.31)	(0.73)	(3.67)	(2.70)